

35-39

OSI 协议  
多媒体通信  
计算机网络

同步通信模型  
同步模型 (9)  
计算机学报 1997 Vol. 27 No. 5

# 一种基于扩展 OSI 协议的多媒体实时同步通信模型\*

A Real-Time and Synchronistic Model of Multimedia Communication Based on Extended OSI Protocol

郑庆华 李人厚

(西安交通大学科学与工程研究院系统工程研究所 西安 710049)

TP39!  
TP393

计算机学报

**摘要** The real-time and synchronization of multimedia communication is the basis of distributed multimedia system. On the basis of the analysis of the problems of multimedia communication, this paper describes the quantized requirements--real-time, isochronism and synchronization, and proposes an extended OSI model which supports multimedia's synchronous communication. Finally, the paper provides a real-time and synchronous model of multimedia communication. This model's implementation have been carried out in the Distributed Multimedia Collaboration System (DMCS).

**关键词** Multimedia, Communication, Synchronization, Real time

多媒体通信不仅要求实时、等时 (isochronism) 地传输多媒体数据,而且要求各种媒体间保持时间和空间上的同步制约。实时性、等时性和同步性是多媒体通信的基本特点,也是实现多媒体通信的关键问题。实时性要求网络能及时传输数据量巨大的连续视频、声音和图象、文本等媒体信息;等时性则要求多媒体数据以稳定的速度均匀、平滑传输,以保持媒体的时基特性;同步性则要求各种媒体在传输后仍然保持原来在时间和空间上的制约关系。

多媒体传输的上述特性不仅要求网络具有足够的带宽或传输率,而且还要求具有支持和保证多媒体同步通信的协议<sup>[2]</sup>。但是,基于传统的 OSI 网络七层协议是面向单一媒体的通信协议,没有提供多种媒体的并发传输、管理和同步等服务。因此,必须扩展 OSI 协议以支持多媒体的实时、等时、同步传输。本文在分析多媒体通信涉及的主要问题之基础上,描述了多媒体同步通信的具体要求,提出了支持多媒体同步通信的扩展 OSI 模型,最后给出了多媒体通信的同步模型。

## 一、同步问题分析

多媒体信息中各媒体间相互依赖、彼此约束。这种约束既存在于空间中,也存在于时间中。多媒体信息的时空约束表明了它在时间和空间上的意义。时

间的意义是指多媒体信息的获取、处理和表现需要一定的时间,并应按照预定的次序平滑地在听觉和视觉中再生;而空间意义则是指多媒体信息在表现时的显示位置和显示形式。时空综合约束反映了多媒体的同步特性、等时特性和实时特性<sup>[1]</sup>。例如 MPEG 声视序列的回放,计算机视频会议中的“面对面”交互等,都体现了多媒体的这些特性。

这里,同步亦是关键,因为媒体间的同步建立在媒体表现时间颗粒化的基础上,同步对时间颗粒度的要求,实际上就是对实时性和粒度均衡性的要求。颗粒度越粗也即同步单位越大,则实时性越差,反之则同步越精确。因此,同步的颗粒度实际上将多媒体的同步性、实时性和等时性统一了起来。

同步颗粒度 (granularity of synchronization) 是一个包含文本、图象、动画、声音、视频等媒体的多媒体对象。将多媒体对象按时间和空间的约束关系有机地合成起来,就是多媒体同步<sup>[3]</sup>。多媒体对象的空间合成是静态的。例如图象显示和文字说明等,而时间合成则是动态的,对合成操作的先后顺序有严格的要求,例如活动影视和它的配音必须一致。但是,空间合成和时间合成又是统一的,因为多媒体对象在整个表现过程中与时间有关,但在表现的每个时间点上又与空间相关。多媒体空间合成和时间合成统称为多媒体合成或多媒体同步。实现多媒体同步

\* 陕西省发展基金项目 (96K07-G5)。郑庆华 博士生,主要研究多媒体技术和 CSCW 技术。李人厚 教授、博士生导师。

的基本思想是：以时间为主线，在时间合成的基础上，附加各个多媒体对象的空间特性（大小、位置等）以实现时空同步。显然，时空同步的关键是时间合成，这就要求相应的同步模型（不论是单机还是分布式）能够充分地描述多媒体对象之间及其内部各媒体间的时序关系，这是同步模型的实质和核心。

目前，实现单机多媒体同步的主要方法有：①基于层次化的同步模型<sup>[2]</sup>；②采用对象（原子对象或多媒体对象）复合的 Petri 网 OCPN (Object Composite Petri-Net) 模型<sup>[3]</sup>描述；③基于参考时间轴 (Reference Timeline) 的同步算法<sup>[4]</sup>；④基于参考点的同步<sup>[5]</sup>。这些方法已经比较有效地解决了单机多媒体同步问题。

但是，在分布式多媒体环境下，由于多媒体信息量的巨大，以及对网络传输的实时性、等时性和同步性的要求，使得多媒体通信变得复杂。目前，多媒体通信的实时性要求可以通过增加网络带宽或传输率、压缩多媒体信息、以及在听觉和视觉允许的范围内，降低视觉和听觉的分辨率等途径解决<sup>[6][7]</sup>。但是，多媒体通信的等时性和同步性要求则只有通过相应的同步协议才能得到保证。这是因为：一方面，目前的 OSI 网络协议是一种面向单一媒体的通信协议，不支持多种媒体的并发传输、管理和同步。另一方面由于网络传输的串行性，以及网络固有的异步特性、传输冲撞和存储设备的延迟等潜在影响，势必造成信息传输的随机延迟，从而破坏了各媒体间的相互约束关系。因此，在多媒体通信中，必须扩展 OSI 网络协议，以支持多媒体同步通信。

从上述分析可以得出：多媒体通信及其同步的关键是时间同步，而时间同步又是在将多媒体数据流颗粒化，并采用相应的同步协议保证多媒体对象传输的等时性和同步性的基础上实现的。

## 二、多媒体同步通信描述

为了便于问题描述，我们先定义以下术语：

**同步时间单位：**媒体间同步的最小时间单位。时间单位决定了同步颗粒度的大小，也决定了同步的精度。取值太大，将影响同步质量，取值太小，又使实现困难。一般地，同步时间单位的大小应以某一基媒体的频率为参考<sup>[2]</sup>。在本模型中，将其定义为播放一帧视频的时间。例如，每秒播放 25 帧，则同步时间单位为 1/25 秒。

**单媒体原子对象 (Atom Object,**

简称 AO)：一个时间单位内所含的单种媒体及其描述信息。单媒体信息能被多媒体硬件或其驱动程序直接识别并支持播放。某一媒体信息流在颗粒化后，形成了一个原子对象队列，如视频媒体的原子对象队列 VAO[1], VAO[2], ..., VAO[n]，声音媒体的原子对象队列 AAO[1], AAO[2], ..., AAO[n] 等等。

**多媒体对象 (Multimedia Object, 简称 MO)：**又称同步颗粒度。是一个时间单位内所含的各种媒体信息及其相互约束关系的集合。多媒体对象可以由若干个单媒体原子对象组合而得到。一个多媒体对象也就是播放一帧的完整信息集。

**投射器对象 (Projector Object, 简称为 PO)：**是一个与多媒体对象或原子对象相关联的、表现为及属性的定义。例如：对原子对象的尺寸、位置、速度、音量等的定义。它实际上是对多媒体对象表现或播放参数的定义。这里，把多媒体数据和表现时所需要的参数（投射器对象）分开，不仅有利于概念的清晰化，而且在实际中还可允许多个不同的多媒体对象使用同一投射器对象。

**多媒体表现：**在投射器对象的作用下，实现多媒体对象或原子对象的表现。对图象和文字而言，表现意味着数据显示；对于音频和视频分量，表现指的是数据动态地在听觉和视觉方面的再生。

多媒体通信同步就是使得经网络传输后的多媒体对象、原子对象和投射器对象序列，仍能保持原来的时空约束关系，在目标结点上仍能得到和源结点相同的多媒体表现。其形式化描述如下：

设时间段  $t_s \sim t_e$  ( $t_s > t_e$ ) 分为  $n$  个时间单位，这里  $n = \lceil (t_s - t_e) / \text{一个时间单位} \rceil$ ，与此相应地有  $n$  个多媒体对象及其相应的投射器对象，分别表示为 MO[1], MO[2], ..., MO[n] 和 PO[1], PO[2], ..., PO[n]，MO[i] 和 PO[i] 的合成在现象上是一个时间单位的多媒体表现，而在信息上则是两者的组合，记为 MPO[i]，即：MPO[i] = {MO[i], PO[i]} ( $i \in [1, n]$ ，下同)。

设  $t_i$  ( $t_i \in [t_s, t_e]$ ,  $t_1 = t_s$ ,  $t_n = t_e$ ) 为源结点上 MPO[i] 的发送时间， $t'_i$  为目标结点上 MPO[i] 的接收时间。则多媒体通信的时间同步可示意为图 1 模型。

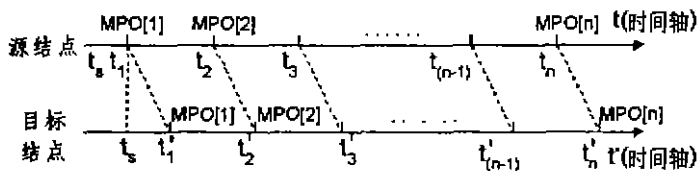


图 1 多媒体通信的时间同步描述

传输  $MPO[i]$  的时间延迟为:  $\Delta T_i = t'_i - t_i$ 。并由此得到:

**定理 1** 实现多媒体传输等时性的充要条件为:任意两个相邻多媒体对象的传输时延相等。即:  $\Delta T_1 = \Delta T_2 = \dots = \Delta T_i \dots = \Delta T_n$ 。

**定理 2** 实现多媒体通信实时性的充要条件为:任意一个多媒体对象的传输时延都小于 1 个同步时间单位。即:  $\Delta T_i < 1$  个同步时间单位,  $i \in [1, n]$ 。

多媒体通信同步则是在保证传输实时性和等时性的基础上,进一步实现多媒体对象的有序发送和接收。即:对  $\forall i, j \in [1, n]$ , 若  $i < j$ , 则应先发送  $MPO$

$[i]$ 。由此可见,多媒体通信同步的实质是:实现多媒体对象经传输后仍保持原来在时基上的配合和制约关系,即统一到原来隶属于同一时间坐标轴上的相对位置。

但是,多媒体通信实际上包括了源结点上多媒体对象合成、多媒体对象传输以及目标结点上多媒体对象接收与解包这三个过程,因此多媒体通信同步应当是从源结点的多媒体对象生成到目标结点的多媒体对象接收与解包为止的端到端同步,其过程如图 2 所示。

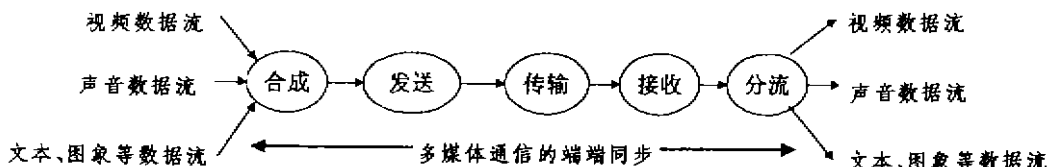


图 2 多媒体通信的端端同步

多媒体通信的端端同步发生在源结点和目标结点上,因此其同步必须由源和目标共同配合实现。在分布式多媒体、CSCW 等系统中,多媒体信息的获取、同步、表现和传输等要由多个过程协同完成,而这些过程分别属于不同层次,这些过程之间需要合作、通信及协调。在结点内部,高层的过程需要调用低层提供的服务;在结点之间,同层之间则要靠相应的协议描述时空同步关系。因此,多媒体通信同步应当采用由源和目标共同控制、分层实施、同层之间对应同步的同步机制。

### 三、OSI 协议的扩展

多媒体通信的物理基础是网络,网络一般采用按协议分层的设计方法。OSI(开放式系统互连模型)是国际标准化组织 ISO 建议的网络协议设计标准<sup>[14]</sup>,已在很多网络操作系统中采用和实现。但是,OSI 参考模型是面向单一媒体的,没有提供多种媒体的多道传送、管理与同步等服务,因此必须对 OSI 参考模型协议进行扩展,使得扩展后的 OSI 协议既能实现单媒体流和多媒体流的传输通信,而且还能满足多媒体通信端端同步的要求。图 3 给出了 OSI 的扩展模型

扩展的 OSI 模型保持底下 5 层不变,在会话层之上,增加了一个逻辑网络驱动层,并在此基础上,将表示层扩展为多媒体获取子层、同步子层和通信接口子层,将应用层扩展为多媒体应用层。其中,逻

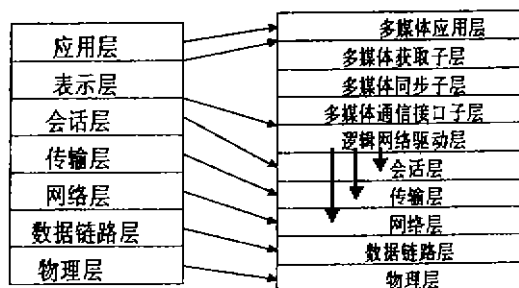


图 3 支持多媒体通信同步的扩展 OSI 模型

辑网络驱动层的主要目的和功能是将各种通信协议如 IPX/SPX、TCP/IP、NETBIOS 等形成一个统一的 API 接口,并实现网络的管理和多媒体信息的指定方式(如广播方式和点对点方式)传输。这样做的好处是使得多媒体通信能基于从网络层到会话层的各种通信协议,多媒体表现及其同步控制则由最上两层实现。

多媒体应用层:实现多媒体表现或应用的用户接口。具体功能包括:①接受用户对多媒体表现的设定。一方面通过界面控制功能如启动、终止、暂停、快进、倒退、继续、切换等实现多媒体表现的控制,另一方面允许用户设定发送的对象、选择需要的媒体及其它交互方式等。②将多媒体表现的要求分解为单机多媒体表现同步和多媒体通信同步。对于单机多媒体表现同步采用 OCPN 模型或基于时间参考轴等算法实现。③对于多媒体通信同步,生成一张多媒

体对象的播放时间表 MPTT (Multimedia Play Time Table), 以记录多媒体对象 MO 间的时序及其相应的表现参数, 并提交至下层。

多媒体表示层: 是分布式多媒体操作系统的核心, 实现多媒体数据流的有序获取、同步表现控制和通信管理。下面以发送结点(源结点)为例说明各子层的功能。接收结点具有类似功能。

多媒体获取子层: ①根据上层的 MPTT, 并以同步时间单位为基准, 并发地获取各单媒体的数据流, 并形成各单媒体原子对象队列, 如视频原子对象队列 VAO[], 声音原子对象队列 AAO[] 等等。这些数据流一方面用于本机回放, 另一方面用于多媒体通信。②根据上层设定的表现参数和媒体间的同步关系, 形成相应的投射器对象队列 PO[]。

多媒体同步子层: ①以同步时间单位为基准, 从各单媒体原子对象队列的队首取出相同序号的原子对象, 并将其合成为多媒体对象后送入 MO 队列中。即:  $MO[i] = \{VAO[i], AAO[i] \dots\}$ 。进一步, 将  $MO[i]$  和  $PO[i]$  合成为  $MPO[i]$ , 并送入 MPO 队列中。②以每隔一个同步时间单位的固定频率向下一

子层提交一个多媒体混合对象 MPO。这样可以保证多媒体对象传输的实时性和等时性。此外, 本子层还将根据网络状态, 实现多媒体同步失调检测, 系统规定了发送相邻两个多媒体对象的最大时间差, 即容忍度, 一般为一个时间单位。当时间差大于容忍度时, 则为同步失调。此时, 将采取强制同步措施实现同步校正。

多媒体通信子层: 从 MPO 队列的队首获取多媒体对象 MPO, 在转换成符合实际网络协议的信包格式后, 调用逻辑网络驱动程序并发送之。

多媒体通信同步的层次结构表明: 多媒体表现与通信交互, 从最上层的宏观同步到最底层的微观细节同步, 是一个自上而下逐层求精的过程。

#### 四、多媒体通信的同步模型

多媒体通信同步需由源结点和目标结点协同控制、分层实施, 并在同步协议的作用下, 实现双方相同层之间的对应同步。这是多媒体通信同步模型的表现实质。根据上述 OSI 扩展层次模型, 设计了如图 4 所示多媒体通信同步模型。

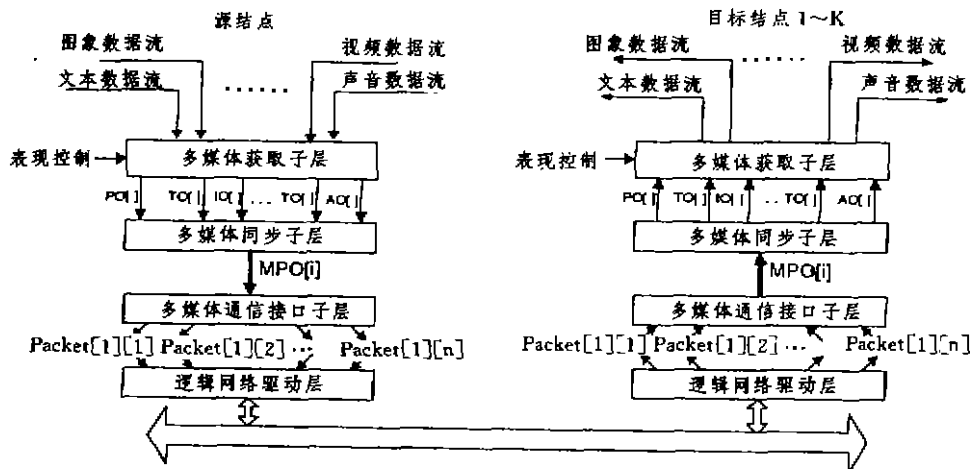


图 4 源结点、目标结点共同控制的同步模型

图 4 示意了多媒体通信及其同步的过程: 从源结点的多媒体数据流获取、原子对象生成(颗粒化)、多媒体对象合成、发送到目标结点的信包接收、多媒体对象的分解直至最后的同步表现。为了实现多媒体对象的等时、实时传输, 应当采用定时发送和接收方式。即以每隔一个同步时间单位的固定频率发送和接收一个多媒体对象。在多媒体对象提交至逻辑网络驱动层后, 进一步将其转换为符合实际网络协议所需的信包格式(如图中的  $Packet[1][n]$ ), 并以更小的时间周期提交给物理网络驱动程序, 这样可进

一步平滑多媒体信息传输。在“分布式多媒体协同操作系统 DMCS”中, 多媒体表示层以每隔 1/25 秒的周期定时向逻辑网络驱动层提交一个多媒体对象, 而逻辑网络驱动层则以 1/125 秒的更小周期发送至网络。不仅平滑了多媒体信息流量, 而且还实现了 25 帧/秒的实时视频传输和回放效果, 超过了 H. 261 规定的 22 帧/秒的传输率<sup>[12]</sup>。

多媒体通信同步是在实时性和等时性的基础上, 进一步保证多媒体对象传输的正确性和有序性。而这主要靠通信协议保证。对声音、视频等动态时基

媒体而言,一方面具有传输数据量大、实时性强的特点,另一方面它们又要求无错传输,但允许有小于2%的传输出错率<sup>[6]</sup>,因此宜采用广播通信方式。对文本、图象及其它系统控制信息等静态媒体,则具有传输数据量较小,实时性要求不高但必须保证无错传输的特点,因此宜采用点对点通信方式。

因为声音和视频采用广播通信方式,因此无法保证传输的正确性。这就要求对多媒体传输进行错误检测,并在传输出错时采取相应的纠错机制使通信能回归正常。为此在源结点上生成多媒体对象和多媒体对象转成信包时,要加上相应的序列号,在目标结点接收时,则根据序列号检测传输的连续性和正确性。一旦发现序列号断序,则认为传输出错,此时应采用纠错机制使通信回归正常。一种简单的处理方法是:目标结点放弃当前多媒体对象的接收,并将上一对象复制给当前对象。值得指出的是:这种纠错方法对声音、视频等动态时基媒体是可以的,但对文本、图象等静态媒体则不适用。

**结束语** 多媒体通信同步是分布式多媒体、计算机支持的协同工作 CSCW 等系统的关键技术之一。本文提出的扩展 OSI 协议,以及由源结点和目标结点共同控制、分层实施、同层之间对应同步的同步模型,不仅很好地解决了多媒体通信的实时性、等时性和同步问题,而且与别的算法<sup>[2][3][6]</sup>相比,本模型还具有不依赖于具体网络结构及通信协议(例如:TCP/IP,SPX/IPX 等)的特点。

本文提出的多媒体实时同步通信模型,已在“分布式多媒体协同工作系统 DMCS”中设计实现,该系统是由西安交通大学研制的基于普通 LAN 的一个 CSCW 应用系统,已于 1996 年 11 月 27 日通过陕西省科委组织的鉴定。实际表明,采用本模型在普通的传输率为 10MB/S 的 LAN 上真正实现了时空上分散的协作成员间的“面对面”实时同步交互,视频传输率达到 25 帧/秒。

## 主要参考文献

- [1] K. H. Smith, Accessing Multimedia Network Service, IEEE Communications Magazine, May 1992
- [2] S. Ramanathan, Feedback Techniques for Intra-Media Continuity and Intra-Media Synchronization in Distributed Multimedia System, The Computer Journal, 36(1), 1993
- [3] Tomas D. C. Little, Network Considerations for Distributed Multimedia Object Composition and Communication, IEEE Network Magazine, Nov. 1990
- [4] B. Gifod etc, On Programming and Supporting Multimedia Object Synchronization, 同[2]
- [5] 李国辉等,多媒体音频视频对象的同步技术研究,计算机研究与发展,32(4):1995
- [6] 钟玉琢等,多媒体计算机技术,清华大学出版社,1993. 5
- [7] 刘红梅等,媒体同步关系描述,计算机研究与发展,33(7):1996
- [8] Z. Tsang-Min Chen etc, Design and Implementation of a Desktop Computer Supported Cooperative Work System, IEEE Trans. on Consumer Electronics, 40(4):1994
- [9] K. Apostolidis etc, A Reservation Protocol for Packet Voice and Data Integration in Unidirectional Bus Networks, IEEE trans. on Communication, 41(3) 1993
- [10] S. Dravida etc, Analysis and Engineering of a Voice/Data Packet Multiplexer, IEEE trans. on Communication, 41(11):1993
- [11] 陈俊华等,多媒体二级同步模型,计算机研究与发展,33(5):1996
- [12] M. Vazirgiannis etc, An Object-Oriented Model for Interactive Multimedia Presentations, 同[2]
- [13] K. Ranai etc, Simulating a Multimedia FDDI Backbone Network, The Computer Journal, 37(1):1994

## 下期内容预告

逻辑程序中的否定问题与非单调逻辑  
遗传优化模糊逻辑控制器  
开放式智能过程中的背景研究需求  
PMS 中的一种匹配通信机制  
KDD: 数据库中的知识发现  
基于曙光并行机的客户/服务器结构的数据开采  
World Wide Web 的索引与查询技术  
数据库系统与 WWW 的集成途径  
基于 Intranet 的集成化信息系统信息的组织和规划