

# 多媒体同步模型研究

李明禄 盛焕焯 孙永强

(上海交通大学计算机科学与工程系, 上海 200030)

**摘要** 多媒体同步模型主要研究多媒体系统中各种媒体对象之间时态关系的规范问题。本文综述了国内外关于多媒体同步模型的研究现状, 包括时间线模型、层次模型及参考点模型, 并指出了进一步的研究方向。

**关键词** 多媒体同步, 时间线模型, 层次模型, 参考点模型

## 1 引言

多媒体研究的先驱者在 80 年代末就开始对多媒体的各个方面进行研究。自 1990 年起形成了多媒体研究的热潮。1990 年 4 月的 IEEE Journal on Selected Areas on Communications 出版了多媒体通信专辑; 11 月在美国 Berkeley 召开了第一届支持数字音频、视频的网络和操作系统国际研讨会 (International Workshop on Network and Operating System Support for Digital Audio and Video), 1991 年、1992 年、1993 年召开的第二届、第三届、第四届同名的国际研讨会的论文集分别被收入 Lecture Notes on Computer Science No. 614、No. 712 和 No. 846。1991 年 10 月 IEEE Computer 出版了多媒体信息系统专辑。1992 年制订了多媒体的两个重要的国际标准 HyTime 和 MPEG。1993 年 ACM 在美国 Anaheim 召开了第一届多媒体年会 ACM Multimedia'93, 并与 Springer Verlag 联合创办了多媒体学报 Multimedia Systems。1994 年 ACM 成立了 SIG Multimedia, IEEE Computer Society 相应成立了 TC on Multimedia, 并于当年 5 月在美国 Boston 召开了第一届多媒体计算与系统年会 (IEEE International Conference on Multimedia Computing and Systems), 重要的国际标准 MHEG 也于当年 8 月制

订。IEEE 的其它期刊和会议、ACM Transactions on Information Systems、The Computer Journal、Computer Communications 以及其它期刊和会议等关于多媒体研究的文章也层出不穷。从此, “多媒体” (multimedia) 这个词不单是商业、市场、政府等用以吸引大众的新名词, 还堂而皇之地登上了世界一流学术界的神圣殿堂<sup>[1]</sup>。

我国的多媒体研究基本上与世界同时起步, 1992 年 12 月在北京召开了第一届全国多媒体技术研讨会。在 1995 年底中国计算机学会成立了多媒体专业委员会。清华大学、东北大学、武汉大学和国防科技大学等都有专门的研究小组从事多媒体的研究与开发。

在上述所有的学术会议、期刊杂志及国际标准中, 关于多媒体同步 (multimedia synchronization) 的研究都是一个重要的课题。到目前为止, 国内研究多媒体同步的文章尚不多见, 而国际上关于多媒体同步的研究则非常活跃 (参见文献 [2~20])。

多媒体对象的时间关系称作多媒体同步, 它是多媒体计算机及多媒体通信领域中的主要问题之一。关于多媒体同步的研究分为: ① 必须规范多媒体对象之间的时间关系; ② 多媒体系统必须保证时间关系的正确表现。本文致力于研究第一个问题, 并对国内外关于多媒体同步模型的研究现状进行了评述, 指明了进一步研究的方向。

## 2 时间线模型

在时间线模型中,所有的媒体对象都与一个时间坐标轴(时间线)相关联,如图 1 所示。这种关系只存在于对象与时间轴之间,所有的对象之间是互相独立的,插入或删除一个对象并不影响其它对象的同步。这种模型的优点是允许对象数据结构的抽象。但是,由于这种模型只能在固定的时间点上定义同步,因此,如果对象没有确定的表现时间(例如,某个对象的表现依赖于用户交互),那么时间线模型将无法处理。

Hodges 等<sup>[2]</sup>在 Athena Muse 项目中首次提出时间线同步模型。通过修改,时间线模型在 Gibbs<sup>[3]</sup>提出的主动对象(active object)模型中得到应用。

Apple 公司推出的 QuickTime<sup>[4]</sup>多媒体标准是 Macintosh 系统软件 System 7.0 的扩展,在 SystemFolder 中加上 QuickTime,就可以象处理图象一样方便地处理多媒体数据。HyTime<sup>[5]</sup>国际标准是一个中性标记语言,用以表示超文本、超媒体、多媒体和基于时间的文献的逻辑结构。在 QuickTime 和 HyTime 中都应用了时间线同步模型。下面我们讨论两个典型的时间线模型:OCPN 和路径表达式。

### 2.1 OCPN

Little 和 Ghafoor 在文献[6]中首次提出了一个基于时间 Petri 网的多媒体同步模型,称作对象合成 Petri 网(OCPN, Object Composition Petri Net)。其实质是对时间线模型的改进。下面给出 OCPN 的定义。

一个 OCPN 定义为一个 6 元组:  $Cocpn = \{T, P, A, D, R, M\}$

这里,  $T = \{t_1, t_2, \dots, t_n\}$ , 表示一个变迁(transition)集合;

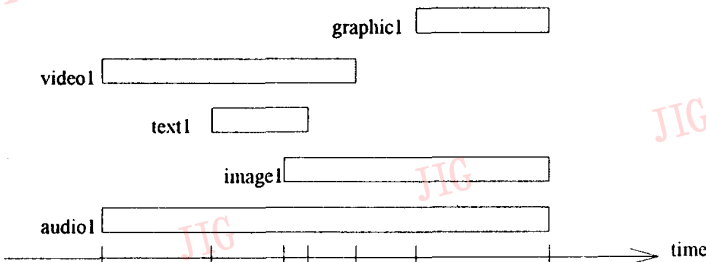


图 1 时间线模型

Fig. 1 Time-line model

$P = \{p_1, p_2, \dots, p_n\}$ , 表示一个库所(place)集合;

$A: \{T \times P\} \cup \{T \times P\} \rightarrow I, I = 1, 2, \dots$ , 表示一个有向弧(arc)集合;

$D: P \rightarrow R(\text{real numbers})$ , 表示一个从库所集合到时段(duration)集的映射;

$R: P \rightarrow r_1, r_2, \dots, r_k$ , 表示一个从库所集合到资源(resource)集的映射;

$M: P \rightarrow I, I = 0, 1, 2, \dots$ , 表示一个从库所集合到标记(marking)集的映射。

与 OCPN 定义相关的是一个说明其语义的点火(firing)规则集。OCPN 的点火规则可总结为:①一个变迁  $t_i$  当它的每个输入库所都包含一个未加锁(unlocked)的托肯(token)时,立即点火;②一旦点火,变迁  $t_i$  删除其每个输入库所的一个托肯,并且在其每个输出库所增加一个托肯;③在接到一个托肯后,一个库所  $p_j$  在给定的时段  $\tau_j$  内保护激活状态,在  $\tau_j$  时段之后,该托肯变为未加锁的。

例 1 用 OCPN 表示一个多媒体的幻灯表现。该表现由一序列同步的具有不同表现时段的音频和视频单元组成。假设有  $n$  个幻灯片要表现,则运用 OCPN 表示如图 2。

通过将 Allen<sup>[7]</sup>的 13 种时片(interval)关系表示成相应的 OCPN 表示,他们提出了通用的 OCPN 模型。根据 OCPN, Ghafoor 领导的研究小组发表了一系列关于多媒体同步的文章。文献[8]增强了 OCPN 的表达能力,阐述了时片的  $N$ -元关系和逆关系。文献[9]提出了扩展的 OCPN, 称作 XOCPN, 不但可以支持多媒体对象的同步,而且还可以支持多媒体对象的通信。XOCPN 增加了对服务质量(QoS)的描述。文献[10]将 XOCPN 应用于预组织(preorchestrated)的多媒体信息的通信。文献<sup>[11]</sup>提出了基于 OCPN 的多媒体信息时空同步的面向对象模型,描述了如何将 OCPN 规范翻译成面向对象语言的算法。Prabhakaran 和 Raghavan<sup>[12]</sup>为 OCPN 增添了人机交互能力。

### 2.2 路径表达式

Hoepner 在文献[13]中引入了一个采用路径表达式的同步模型,路径表达式是创建时态结构的早期代数之一。为了描述多媒体对象表现的同步, Hoepner 将 Allen 的 13 种时片关系表示成路径表达式,并选用了如下的路径操作符集合,这里  $A$  和  $B$  是多媒体表

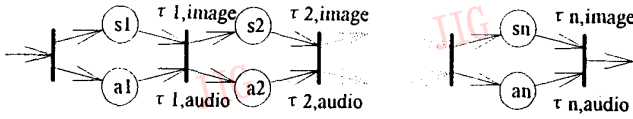


图 2 多媒体幻灯表现的 OCPN 表示

Fig. 2 OCPN representation for multimedia slide presentation

现动作,而  $N$  是一个正自然数。路径操作符的名称、含义按优先权增序排列如表 1 所示。

表 1 路径操作符

Table 1 Path operators

路径操作符	名称	含义
$A \wedge B$	Parallel-last	动作 $A$ 和 $B$ 同时开始并行执行,复合动作 $A \wedge B$ 当所有参与的动作终止时终止。
$A \vee B$	Parallel-first	动作 $A$ 和 $B$ 同时开始并行执行,复合动作 $A \vee B$ 在第一个(时间上)参与的动作终止时即终止。
$A ; B$	Sequential	只有当 $A$ 执行后才能执行 $B$ , $A$ 的结束点等于 $B$ 的开始点,复合动作 $A ; B$ 当序列中的最后一个动作终止时才终止。
$A   B$	Selective	或者执行 $A$ 或者执行 $B$ , 选择依赖于某一条件,该条件并非路径表达式的一部分,但它能被别的实例检测到,复合动作 $A   B$ 在被选中的那个动作终止时即终止。
$A^*$	Repetition	动作 $A$ 重复执行 $i$ 次。若 $i$ 省略,则 $A$ 将执行 0 次或许多次,精确执行的重复次数依赖于别的实例。
$N ; A$	Concurrency	动作 $A$ 并发执行 $N$ 次。如果 $N=1$ ,则 $A$ 的执行是独立的。若 $N=infinite$ ,则将不允许 $A$ 的并发执行。缺省值为 $N=1$ 。

例 2 一个路径表达式描述:  $path A ; ((B \wedge C) DV (D \wedge E)) ; F^* end$

其含义为:动作  $A$  开始执行。精确地,在  $A$  的结束点。四个动作  $B, C, D$  和  $E$  同时开始执行。一旦  $B$  和  $C$  或者一旦  $D$  和  $E$  终止,动作  $F$  开始执行。  $F$  可执行 0 次或多次,整个路径只允许执行一次。

### 3 层次模型

在层次同步模型中,采用两个主要的同步操作符:动作(action)间的顺序同步和并行同步。一个动作可以是原子动作或是组合动作。原子动作与某一媒体对象的表现相关联;而组合动作由原子动作与联结它们的同步操作符的集合组成。

如图 3 所示,运用层次同步模型的多媒体表现可看作是一棵树,其叶子是原子动作(或表现对象),而其节点是同步操作符。组合动作可以看作是树的某一个子树。层次同步模型的主要缺点是同步只允许发生在动作的开始点或结束点。

在文献[14]中首次提出了层次同步模型。Shepherd 和 Salmony<sup>[15]</sup>运用层次同步模型扩展了 OSI 参考模型以支持多媒体应用。Amsterdam 超媒体模型<sup>[16]</sup>也采用了类似的同步方法。

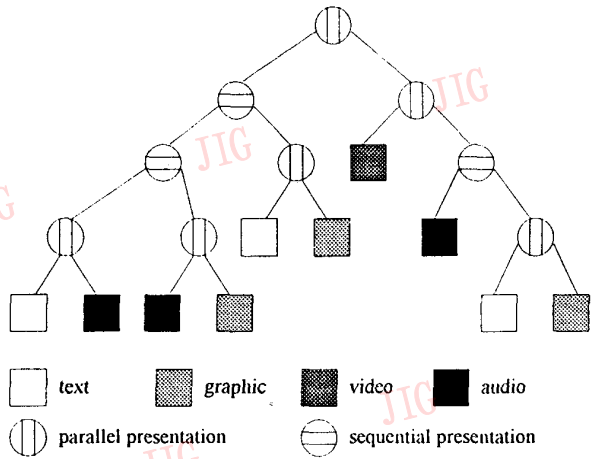


图 3 层次同步模型

Fig. 3 Hierarchical synchronization model

### 4 参考点模型

Steinmetz 在他的关于多媒体同步性质研究的著名文章<sup>[17]</sup>中首次提出参考点同步模型。这种模型的思想源于 Hoare 的 CSP。

在这种模型中,时间相关的(time-dependent)多媒体对象如视频或音频被看作是由时间无关的(time-independent)在恒定的时间间隔中表现的子单元(例如视频的帧,音频的样(sample))序列组成。同步被规范为将在同时表现的不同对象中的子单元

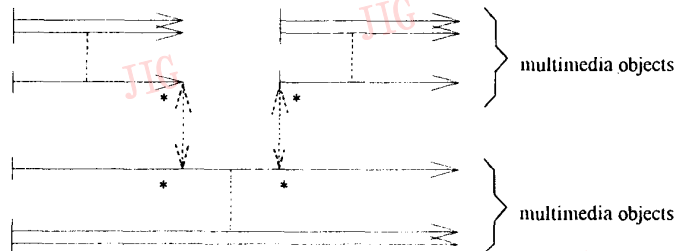


图 4 参考点同步模型

Fig. 4 Reference point synchronization model

之间的关系,如图 4 所示。

### 4.1 基于 CCS 的同步模型

Eun 等<sup>[19]</sup>提出了一个基于 Milner 的 CCS 的多媒体同步模型,CCS 是一个描述代理(agent)系统的演算,该系统可通过消息同步或异步通信来执行。一个由 CCS 描述的系统可以看作是由代理及连接它的散布消息的通道组成的网络。类似地,一个多媒体系统可以看作是由多媒体对象互相连接而成的网络,每个对象对应于一个代理,它们之间通过交换消息来实现同步。这与基于 CSP 的参考点模型在原理上是相通的。

基于 CCS 的多媒体同步的规范机制可以描述如下,每个对象的输入端口记作  $a?$ , 输出端口记作  $a!$ 。如果一个对象 A 有一个输入端口  $a?$  和输出端口  $b!$ , 可以写成  $A = a?. b!$ . A, 其语义为该对象一旦接收到一个来自输入端口  $a?$  的消息就开始执行表现,而在表现结束时通过输出端口  $b!$  散布一个消息,这与 CCS 的语义是相同的。

例 3 (参见图 5)在图 5(a)中,CCS 表达式的含义为两个对象 A 和 B 顺序同步,对象 B 在对象 A 完成表现后才开始表现。在图 5(b)中给出一个较复杂的例子。两个对象 A 和 B 同步地同时开始和同时结束表现。在图 5(c)中给出了一个类似于唇同步(lip-synchronization)的例子。对象 Video 指一个运动视频,而 A, B 和 C 指精确对应于运动视频某个时间位置上的声音。当 Video 对象接收到一个 Start 消息  $s$  时,它开始了一段没有声音的视频表现。如果它到达第一个设计好的时点,它就发出一个 Start 消息  $s1$  给声音对象 A, 然后 A 将与 Video 一起并行地表现。其它声音对象的表现可以此类推。

## 5 结束语

Meira 和 Moura 在文献[20]中评价了时间线模型、参考点模型和层次模型的优缺点,认为这三类模型都不能满足多媒体同步规范的需要。

我们认为,Allen 的 13 种时片关系不能满足多媒体同步规范的需要,因为在多媒体系统中存在大量的用户交互及许多不可预知的表现时段。由于同步主要研究多媒体对象的时态关系,因此从理论上探讨究竟存在多少和哪些时态关系这个问题非常必要,在此基础上,有望开发出能规范多媒体系统中所有时态关系的多媒体同步通用模型<sup>[21]</sup>。

### 参考文献

- 1 李明禄. 多媒体同步:通用模型、描述语言及其实现:[博士学位论文]. 上海:上海交通大学,1996.
- 2 Hodges M E et al. Athena Muse:a construction set for multimedia applications. IEEE Software,1989;37~43.
- 3 Gibbs S. Composite multimedia and active objects. Proc. of OOPSLA'91,1991.
- 4 Hoffert E et al. Quick Time:an extensible standard for digital multimedia. COMCON Spring'92. San Francisco,CA,1992. 2,15~20
- 5 ISO, Hypermedia/time - based document structuring language (HyTime), ISO/IEC DIS10744,1992.
- 6 Little T D C,Ghafoor A. Synchronization and storage models for multimedia objects. IEEE Journal on Selected Areas in Communications,1990. 4,8(3):413~422.
- 7 Allen J F. Maintaining knowledge about temporal intervals. Communications of the ACM,1983,11,26(11):832~843
- 8 Little T D C,Ghafoor A. Interval-based conceptual models for time-dependent multimedia data. IEEE Transactions on Knowledge and Data Engineering,1993. 8,5(4):551~563
- 9 Qazi N,Woo M,Ghafoor A. A synchronization and communication model for distributed multimedia objects. Proceedings of ACM Multimedia'93,USA: Anaheim,1993;147~155.
- 10 Woo M,Qazi N U,Ghafoor A. A synchronization framework for communication of preorchestrated multimedia information, IEEE Network, 1994, 8(1):52~61.
- 11 Iino M,Day Y F, and Ghafoor A. An object-oriented model for spatio-temporal synchronization of multimedia information. In: 1st IEEE International Conference on Multimedia Computing and Systems. Boston,MA,1995. 5:110~119.
- 12 Prabhakaran B, Raghavan S V. Synchronization models for multimedia presentation with user

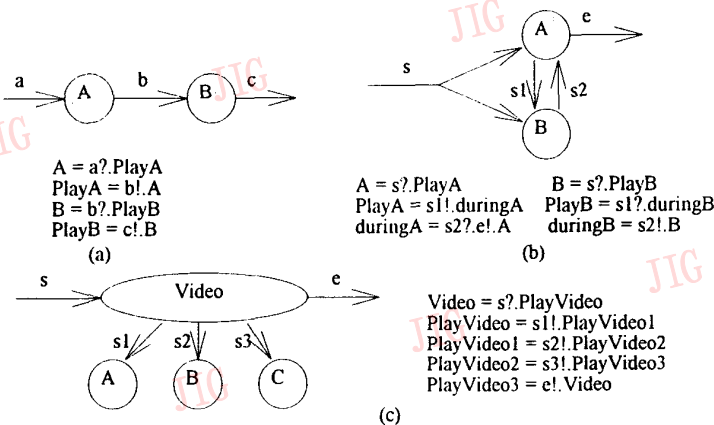


图 5 运用 CCS 规范多媒体应用

Fig. 5 Specify multimedia applications using CCS

- participation, Proceedings of ACM Multimedia'93, USA; Anaheim, 1993. 8:157~166.
- 13 Hoepner p. Synchronizing the presentation of multimedia objects. Computer Communications, 1992. 11, 15(9):557~564.
  - 14 ISO IE JTC 1/SC 18/WG3, Multimedia Synchronization; Definitions and Model. Input Contribution on Time Variant Aspects and Synchronization in ODA-Extension, 1989. 2.
  - 15 Shepherd D, Salmony. Extending OSI to support synchronization required by multimedia applications. Computer Communications, 1990. 9, 13(7):399~406.
  - 16 Von Rossum G, Jansen J, Mullender K S et al. CMIFed; a presentation environment for portable hypermedia documents. Proceedings of ACM Multimedia'93, USA; Anaheim, 1993. 8:183~188.
  - 17 Steinmetz R. Synchronization properties in multimedia systems. IEEE Journal on Selected Areas in Communications, 1990. 4, 8(3):401~412.
  - 18 Blakowski G, Hubel J, Langrehr et al. Tool support for the synchronization and presentation of distributed multimedia. Computer Communications, 1992. 12, 15(10):611~618.
  - 19 Eun S B et al. Specification of multimedia composition and a visual programming environment. Proceedings of ACM Multimedia'93, USA; Anaheim, 1993. 8:167~173
  - 20 Meira S R L, Moura A E L. A scripting language for multimedia presentations. In: 1st IEEE International Conference on Multimedia Computing and Systems, Boston, MA. 1994. 5:484~489.
  - 21 Li Minglu, Sun Yongqiang, and Sheng Huanye. Generalized synchronization model and description language for multimedia systems. In: 3rd IEEE International Conference on Multimedia Computing and Systems, Japan; Hiroshima, 1996. 6.



李明禄, 博士, 副教授, 1985年毕业于解放军电子技术学院, 1993年获上海交通大学博士学位, 现为上海交通大学博士后。主要研究领域为多媒体计算机技术、并行与分布系统和高速信息网络。

## On Multimedia Synchronization Models

Li Minglu, Sheng Huanye, Sun Yongqiang

(Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200030)

**Abstract** The main issue in the reserach of multimedia synchronization models is specification of the temporal relations in multimedia objects. This paper reviews the state of the art in multimedia synchronization models, including time-line models, hierarchical models and reference point models, and points out the further research directions.

**Keywords** Multimedia synchronization, Time-line models, Hierarchical models, Reference point models