

# 基于 MPEG-4 的多媒体同步模型设计

张爱军, 王 石

(苏州大学 计算机科学与技术学院, 江苏 苏州 215006)

**摘要:**文中旨在研究基于 MPEG-4 技术的面向对象多媒体同步。所采用的方法是在分析现有多媒体同步模型的基础上,结合 MPEG-4 系统同步机制和高级同步模型(FlexTime),设计面向对象的多媒体同步模型,以实现 MPEG-4 中各媒体对象间同步。由此方法设计的层结构同步模型是一种有效可行的多媒体间同步的实现方法。和其它的多媒体间同步方法相比,这种方法在基于对象的多媒体的传输过程中携带较少的时间信息来完成对象间的同步,并且增强了基于对象内容交互的灵活性。

**关键词:**同步;MPEG-4;高级同步模型;层结构同步模型

**中图分类号:**TN919.8

**文献标识码:**A

**文章编号:**1005-3751(2006)04-0049-03

## A Multimedia Synchronization Model Based on MPEG-4

ZHANG Ai-jun, WANG Shi

(Computer Science and Technology School of Soochow University, Suzhou 215006, China)

**Abstract:** The paper focuses on multimedia synchronization based on MPEG-4. By analyzing existing multimedia synchronization model, the paper designs the object-oriented model to synchronize all MPEG-4 multimedia objects, based on MPEG-4 system synchronized mechanism and the Advanced Synchronization Model (FlexTime). The proposed layered structure synchronization model is an effective and feasible method of multimedia synchronization. Compared with other methods, the proposed one uses less time information to finish object synchronization during the object-oriented multimedia transmission and improves the flexibility of the object-oriented content interaction.

**Key words:** synchronization; MPEG-4; FlexTime model; layered structure synchronization model

### 0 引言

多媒体同步是一个涉及领域广泛而且处理过程复杂的问题。多媒体同步由很多系统部件来支撑和体现,包括操作系统、通信系统、数据库、文档,甚至一些具体应用,因此,多媒体系统中的同步必须在多个层次予以考虑。

### 1 分布式环境中多媒体同步的概述

多媒体同步可分为两大类:媒体内同步(intra-synchronization)和媒体间同步(inter-synchronization)。前者是指媒体在播放时要求保持媒体内部的连续性,如25帧/秒播放TV,要求画面的连续性;后者是指要保持媒体间的同步时间关系,如视/音频数据的播放同步<sup>[1]</sup>。

在网络环境中,影响多媒体同步的重要因素是延迟、抖动和时钟差异。这些问题是密切相关的,不能单独进行解决,必须综合考虑。MPEG-4系统中的对象内容包括

时基媒体和非时基媒体两种,如视音频的连续时序数据流需要以时间为参照系进行有序的组织,属于时基媒体,而文本、静态图像插入到播放流对时间的要求不是很严格,属于非时基媒体。组织时基的、非时基的多种媒体序列以达到某种表现效果的任务,就是实时多媒体的同步。同步技术既可用于并发或顺序的数据流布局,也可用于对所产生的外部事件管理。

### 2 同步技术的形式化描述

利用形式化描述方法可以对多媒体同步关系进行各种描述。由于同步关系是媒体时间关系的描述,一般需要将时间关系(包括时间的定义)引入到描述中。

同步技术的形式化描述分为下面几类<sup>[2]</sup>:

#### (1) 基于事件的描述。

在这种描述方法中,媒体的表现行为由事件触发,典型的表现行为包括表现的开始、表现的停止、表现的准备。触发表现行为的时间是由定时器生成的外部行为,或是由于时间相关媒体符合一定条件而产生的内部行为。

#### (2) 基于 Petri 网的描述方法。

Petri 网是一种并发系统及其同步关系的建模工具。简单 Petri 网的转移触发时间是不确定的,转移触发为一

收稿日期:2005-08-14

基金项目:江苏省自然科学基金资助项目(BK2003029)

作者简介:张爱军(1975—),男,山西阳泉人,硕士,研究方向为多媒体技术、网络与数据库;陆建德,教授,研究方向为网络协议、网络内核分析与设计、网络安全。

即时事件。为了描述 Petri 网中的非零时间费用,需要在 Petri 网中引入时间约束,这种增强型 Petri 网称为定时 Petri 网(TPN)。以定时 Petri 网为基础,可以描述多媒体信息的同步关系。

### (3) 基于时间间隔的描述。

这种描述将表示对象的表现持续时间作为时间间隔,两个时间间隔之间将采用 7 种不同类型的时间关系表达对象之间的同步逻辑,设 A、B 各代表一个媒体流,则它们间相互关系可表示为:A 在 B 前;A 在 B 后;A 与 B 重叠;A 在 B 内部;A 与 B 同时开始;A 与 B 同时结束;A 与 B 相等。

## 3 MPEG-4 系统同步机制

### 3.1 相关概念

#### 3.1.1 基本码流(Elementary Stream, ES)

系统将从压缩层得到的码流数据作为 ES 数据。ES 在服务器端的压缩层产生,通过同步层、传输层发送给客户端,再由客户端的传输层、同步层传递给压缩层。这个过程是一个端到端的发送,从而保证了 ES 的完整性。

ES 并不是最小的处理单位,它包含以下更加详细的组成结构:

#### (1) AU(Access Unit, 访问单元)。

ES 被划分成多个 AU。AU 是 ES 数据中具有时间属性的最小实体。AU 的内部组成结构由产生基本码流的压缩层定义, AU 由 ES 中所有具有相同处理时刻的合成单元构成,同一个 ES 中没有两个 AU 具有相同的时刻。

#### (2) CU(Composition Unit, 合成单元)<sup>[3]</sup>。

AU 在接收端压缩层解码以后得到 CU。CU 对于 ES 来说是不可见的, CU 的基本属性包括 CTS(合成时间戳)、生命周期和标签等。CU 是组成 AU 的最小实体,单个 CU 表示单个 IPMP 消息、OD 命令、BIFS 命令帧/BIFS 动画帧、视频对象数据记录、STB 单元记录或者 OTB 单元记录等基本实体。

#### 3.1.2 同步信息

MPEG-4 系统中对象的同步是通过以下 4 个时间信息实现的:

(1) 系统时基(STB):系统时基定义了终端的时间戳,它是各种时间戳的时间参考。

(2) 对象时基(OTB):对象时基定义了一个对象单元的时间戳。

(3) 对象时钟参考(OCR):用来传输 OTB 给基本码流解码器。

(4) 解码时间戳(DTS):每个访问单元都有一个表示解码时间的标志,即在该解码时间,这个访问单元必须能被用于解码,在该解码时间之前,访问单元不能使用。

(5) 合成时间戳(CTS):每个 CU 都有一个相关联的表示合成时间的时间戳,它表示在该时刻在合成存储区内该 CU 可以被用来合成,在该时刻之前 CU 不能被处

理<sup>[4]</sup>。

### 3.2 同步层

同步层指定 ES(基本码流)数据以 AU 或 AU 一部分打包的语法,经过同步层得到的数据包称为 SL 数据包。SL 数据包包含的内容只能是单个 AU 或者单个 AU 的一部分。一方面,发送端 SL 将 ES 和同步信息封装并映射成一个同步数据流包序列(SPS),再将它们传输到传输层,这些数据包不仅含有定时、分段和时间戳信息,而且还有随机访问信息。另一方面,接收端 SL 从传输层接收 SPS,从同步数据包中提取同步信息,再根据同步信息重新生成完整的 ES,为 AU 解码和解码以后的合成做准备。一个 SL 数据包不包含它的长度指示,而是由处理同步包分组数据流的传输层定义相关的语法和语义,所以 SL 数据包必须用一个合适的传输协议来发送和接收(如 RTP 传输协议)。同步层是 MPEG-4 中实现定时和同步的重要机制。

## 4 高级同步(FlexTime)模型

### 4.1 FlexTime 模式

高级同步模型<sup>[5]</sup>增强了普通 MPEG-4 定时模式的功能,允许多种流和对象的同步,这些流或对象可能是视频、音频、文本、图像,甚至是程序,并且可能从多个数据源产生。

普通 MPEG-4 定时模式主要为“推”式广播应用而设计,在这些应用中,存取单元间的时间同步是通过“硬”时间戳和参考时钟得到的。这种机制在提供精确的流间同步的同时,却难于对从不同源(并可能具有不同的参考时钟)来的流提供流内同步,而这种情况在绝大多数因特网应用和复杂的广播应用中常常遇到。

### 4.2 灵活的时期

在一个不可靠交付的环境里,一个基本码流或流的一部分,可能比要求的时间晚些交付。FlexTime 模式基于一种比喻为“弹簧(spring)”的方法,使其对交付延迟不敏感。一个弹簧具有 3 个常数:最小长度(小于此值,弹簧不能再压缩),最大长度(超过此值,弹簧将损坏),和最佳长度(在此长度时,弹簧正常地静止)。

### 4.3 相对的开始和结束时间

通过定义相同的时间开始(“CoStart”),或相同的时间结束(“CoEnd”),或者使一个的结束时间与另一个的开始时间相合(“Meet”),可以使两个或更多个基本码流或流片段达到彼此同步。

### 4.4 MPEG-4 中对 FlexTime 的支持

MPEG-4 中有两个节点和一个描述符支持 FlexTime。这两个节点是 TemporalTransform 节点和 TemporalGroup 节点,还有一个描述符 SegmentDescriptor。TemporalTransform 节点指定一个需要同步(或灵活安排)的 MPEG-4 对象的时域特性。TemporalGroup 节点指定由 TemporalTransform 节点标识的对象间的时域关系,而

SegmentDescriptor 标识一个流的可以被同步的部分。

可以实现分布式多媒体的同步。

## 5 层结构同步模型的设计

层结构同步模型是用对象逻辑关系、时间戳和系统时基显示或隐式地描述各个对象的具体时间信息(如对象生成时间戳,对象合成时间戳),以达到用最少的基于时间轴的时间戳来描述各个对象用于同步的时间信息,从而使得基于对象的多媒体在传输过程中携带较少的时间信息完成对象间的同步以及增强了基于对象内容交互的灵活性。

层结构同步模型如图 1 所示,该图表示了系统时基层下各对象之间的时序关系和同步关系。该模型分为 4 层,按照从上到下的结构分别为系统时基层、时基相关层、定时相关层和瞬时相关层。系统时基层表示作为时间系统轴的系统时基组,单个系统中各媒体的系统时基是相同的,分布式多媒体由于存在时间差异,各媒体可能有不同的时基;时基相关层表示具有同一系统时基的对象内容组成的多个以不同时间戳作为参考的数据流,图 1 中 ES11 和 ES12 代表了两种具有不同时间戳的时基数据流;定时相关层表示时基相关层中每一个时基数据流所包含的对象内容组,一个时基数据流中的多个对象内容组之间以默认不变的时间间隔区分,也就是说时基数据流中连续的对象内容组存在一个固定的时间差,如图 1 中对对象内容组 AU111 和 AU112 之间、AU121 和 AU122 之间、AU122 和 AU123 之间存在着默认的时间间隔;瞬时相关层则表示定时相关层的中每一个对象内容组包含的对象内容具有相同的开始时间,如对象内容 CU1111, CU1112 和 CU1113 具有相同时刻的开始时间。

在该模型下,每个对象内容赋予一个唯一的四维下标,四维下标的数值分别对应于系统时基层、时基相关层、定时相关层和瞬时相关层的位置。

使用层结构的同步模型可以在瞬时相关层通过逻辑关系尽最大可能地组合各种对象内容,从而实现了用最少的基于时间轴的时间戳来描述各个对象的时间信息。

层结构的同步模型还提供了灵活地修改对象内容时间信息和重新进行对象同步的能力。如图 2 所示,通过显式地修改 ES11 的时间戳,从而隐式地修改 ES11 内所有包含的对象内容的时间信息。

层结构同步模型的最高层为系统时基层,每一 ES<sub>ij</sub> 中包含的对象内容隐式地对应于该对象内容所属的系统时基数值,当不同系统时基层内的对象进行事件交互时,

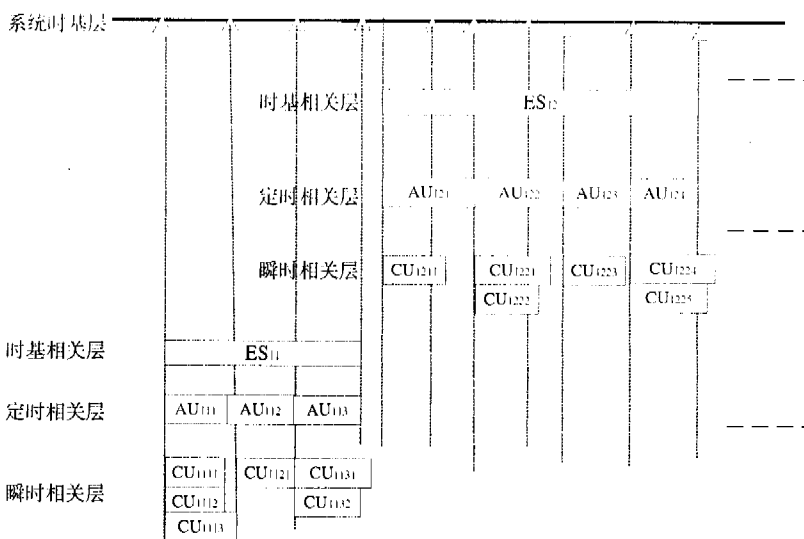


图 1 层结构同步模型示例(一)

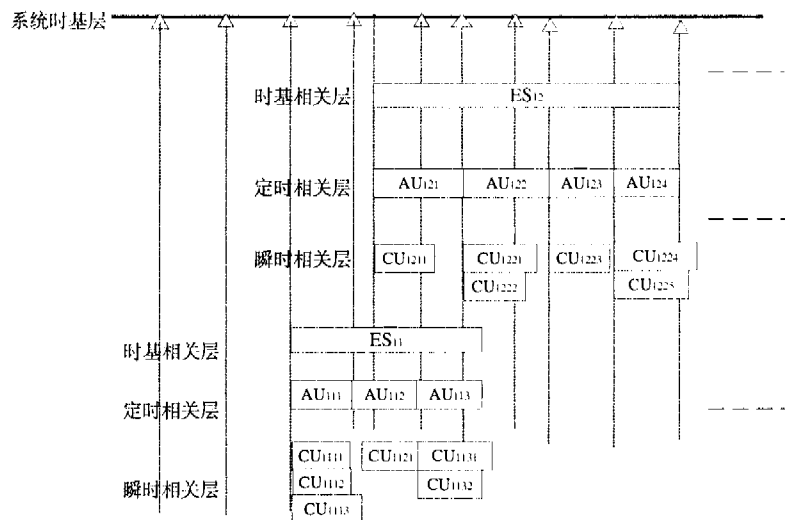


图 2 层结构同步模型示例(二)

视音频实时传输原型系统采用层结构的同步模型对各对象同步信息进行处理,减少传输数据量,并提高了基于对象内容的交互能力。

## 6 结 论

终端多媒体同步是保证多媒体演播质量的关键因素。现有的方法主要是通过时间参考、时间戳和接收缓冲区来调整同步质量。文中在现有的同步技术基础上,基于 MPEG-4 技术中面向对象的同步特点,设计了面向对象的层结构的同步模型,可用于解决数据传输过程中单媒体内部和分布式多媒体之间的同步问题。

## 参考文献:

- [1] 孙文彦. 实时传输中的多媒体同步技术研究[D]. 北京:北京航空航天大学, 2001.
- [2] Blakowski G, Steinmetz R. A Media Synchronization Survey: Reference Model, Specification, and Case Studies[J]. IEEE

(下转第 54 页)

灰类即为该指标所属类型<sup>[4]</sup>。

(8) 对评估结果进行分析。

## 2 软件需求风险的综合评估

假设有一个软件,共有 4 个需求(A,B,C,D)要进行风险评估,所需数据见表 1。

表 1 需求风险评估实验数据

单元	风险	项目 1 性能	项目 2 安全	项目 3 过程	项目 4 技术	项目 5 数据库	项目 6 日程	项目 7 外部	项目 8 稳定
单元 1	A	0.33	0.29	0.10	0.22	0.17	0.12	0.23	0.12
单元 2	B	0.41	0.15	0.32	0.16	0.25	0.26	0.16	0.28
单元 3	C	0.32	0.28	0.16	0.25	0.19	0.14	0.29	0.11
单元 4	D	0.36	0.24	0.18	0.26	0.11	0.19	0.24	0.17

### 2.1 灰聚类评估指标体系

①聚类单元。由表 1 可以知道,4 个聚类单元分别为需求 A,B,C,D。

②聚类项目。由表 1 可以知道 8 个聚类项目分别为性能、安全、过程、技术、数据库、日程、外部和稳定。

③灰类。共有 4 个灰类,分别是:

$k = 1 \Rightarrow$  第 1 灰类“较大”     $k = 2 \Rightarrow$  第 2 灰类“大”  
 $k = 3 \Rightarrow$  第 3 灰类“小”         $k = 4 \Rightarrow$  第 4 灰类“较小”

### 2.2 构造样本矩阵

令  $d_{ij}$  表示  $i$  单元对于  $j$  项目的样本,基于表 1 建立样本矩阵  $d$ 。

### 2.3 确定各指标极性及其界限

因为在软件开发过程中,各指标的数值越小,需求的灰度越小。所以各指标的极性相同,均为正极性指标。由于各指标的涵义不同,量纲也不同,故不能确定同一个界限,而应根据指标的意义,参考各指标数据的分布特征,确定各指标<sup>[5]</sup>。各灰类界限值分别列于表 2。

表 2 各指标灰类的界限

类别	性能	安全	过程	技术	数据库	日程	外部	稳定
较大	0.50	0.45	0.60	0.70	0.80	0.60	0.80	0.90
大	0.40	0.35	0.50	0.50	0.60	0.50	0.60	0.70
小	0.30	0.20	0.20	0.30	0.40	0.30	0.30	0.50
较小	0.10	0.10	0.10	0.20	0.20	0.20	0.10	0.30

### 2.4 确定白化权函数并计算综合权系数矩阵

根据表 2 的类别界限,各指标的白化权函数及权系数的计算公式见上文。根据各指标对需求风险的影响程度,赋予各指标的权重为:

$$A = [p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8] \quad (8)$$

### 2.5 计算折算系数和评估值

根据上面的公式,得到各个需求风险的折算系数矩阵

$\eta_{ij}$ 。

各指标的权重为  $A = [0.2, 0.2, 0.1, 0.05, 0.05, 0.1, 0.1, 0.2]$ ,

根据上面的公式,对于需求 A 来说,得到

较大      大      小      较小

$$\sigma_1 = (0, \quad 0.0566, \quad 0.06112, \quad 0.0563)$$

$$\sigma_{1k} = \max(0, 0.0566, 0.0611, 0.0563)$$

$$= 0.0611 = \sigma_{13}$$

表明  $k^* = 3, (i = 1) \in \{k^* = 3\}$ , 即需求 A 的风险等级为灰类 3:“小”。

在得出需求风险等级的结论以后,不应当简单地将其置之一边,而应进一步进行分析。因为软件需求风险分析的目的,不仅是要得出需求是否具有风险的结论,而且要通过对其结果的分析,找出影响软件需求的主要因素,通过改进方案或采取其它措施,降低软件需求的风险。

## 3 结束语

文中尝试将灰色理论的灰色聚类评估理论应用于软件需求风险分析,提供了一种量化的分析软件需求风险的方法。但是,鉴于灰色理论自身的特点,其正确性要依赖于各种自定义因子的正确性。因此,对综合评估理论中单元集、项目集、评估矩阵、各项目权值、结论模型等的定义将直接影响最后的结论<sup>[6]</sup>。根据各类应用软件需求的不同特点和对软件需求分析案例的统计分析,进一步量化各种因素对需求风险的影响,对于软件需求风险分析的研究具有更深远的意义。

### 参考文献:

- [1] 邓聚龙. 灰理论基础[M]. 武汉:华中科技大学出版社, 2002.
- [2] Sommerville I, Sawyer P. 需求工程[M]. 北京:机械工业出版社, 2003.
- [3] 王学萌,张继忠,王 荣. 灰色系统及实用计算程序[M]. 武汉:华中科技大学出版社, 2001.
- [4] 罗佑新,熊国华,曾经梁,等. 计算机软件质量综合评价的灰色聚类模型及应用[J]. 计算机工程与应用, 2003, 39(8):84-85.
- [5] 冯建湘,唐 嵘,王双维. 软件质量要素的灰色关联分析及其应用[J]. 计算机工程, 2004, 30(9):91-92.
- [6] Chen Pei-Yin, Jou Jer Min. Adaptive arithmetic coding using fuzzy reasoning and grey prediction[J]. Fuzzy Sets and Systems, 2000(114):239-254.

(上接第 51 页)

Journal on Selected Areas in Communications, 1996, 14(1):5-35.

- [3] ISO/IEC 14496-1 International Standard, Coding of Audio-Visual Objects: Systems[S]. 2001-12.

- [4] 钟玉琢,王 琪,贺玉文. 基于对象的多媒体数据压缩编码国际标准 MPEG-4 及其校验模型[M]. 北京:科学出版社, 2000.

- [5] ISO/IEC JTC1/SC29/WG11 N3747, Overview of the MPEG-4 Standard[S]. 2000-10.