

音视频同步技术综述^①

柴若楠, 曾文献, 张鹏云

(河北经贸大学 信息技术学院, 石家庄 050061)

摘要: 介绍了多媒体数据的构成、多媒体同步的分类和多媒体同步的定义, 在此基础上介绍了多媒体中音视频同步的相关技术, 并对不同技术的优缺点进行了阐述。最后, 介绍了目前较为流行的音频嵌入视频技术的应用。

关键词: 多媒体通信; 音视频同步; 时间戳同步; 嵌入式音视频同步

Survey on Audio and Video Synchronization

CHAI Ruo-Nan, ZENG Wen-Xian, ZHANG Peng-Yun

(School of Computer Science and Technology, Hebei University of Economics and Business, Shijiazhuang 050061, China)

Abstract: This paper describes the composition of multimedia, the classification and definition of multimedia synchronization at first. Then, it introduces the related technology of audio and video synchronization and their advantages and disadvantages. At last, it introduces the application that uses the embedded audio and video synchronization which is the most popular technology.

Key words: multimedia communication; audio and video synchronization; timestamp synchronization; embedded audio and video synchronization

1 引言

音视频同步是多媒体系统服务质量(QoS Quality of Service)研究中的一项重要内容。在视频会议、可视电话、视频点播等多媒体应用中, 由于网络传输中的延迟、抖动、时间偏差、网络传输条件的变化以及发送端的发送速度与接收端的接收速度的不匹配等问题, 使得接收端的媒体存在异步现象。通过媒体同步技术可解决这些问题, 其中音视频的同步是一个关键性技术。

目前为止, 国内外已提出多种同步模型及同步方案。例如: Escobar 等人提出适用于多种通信模式的流同步协议是需要全网同步时钟的自适应同步方案^[1]; Rangan 等人提出一种适用于多个信源一个信宿的基于反馈的同步技术^[2]; 基于实时通信协议的同步方案^[3,4]; 利用语音与其唇形之间的对应关系的同步方案^[5]以及音频嵌入视频同步方法等。

2 多媒体同步的基本概念

2.1 多媒体数据的分类

由在内容上互相关联的文本、数据、图形、图像、动画、音频和视频等媒体数据构成的一种复合信息实体即多媒体数据。根据媒体数据与时间之间的相互关系, 可以将媒体数据划分为静态媒体数据和连续媒体数据。按照数据对时间的关系和数据生成方式的差别, 可以将不同媒体类型的数据划分为表 1 所示。

表 1 媒体数据的成分^[6]

分类	连续媒体	静态媒体
源自现实世界	语音视频信号	静态图像
由计算机完成	动画	文本、图形

在多媒体系统中, 通常从多媒体的不同侧面来表达媒体之间存在的相互依存的关系。下面分别从三个不同的角度对媒体同步进行分类^[6]。如图 1 所示。

① 基金项目:河北省应用基础研究计划重点基础研究项目(10963529D);河北省高等学校科学技术研究重点项目(ZD2010213)

收稿时间:2011-03-21;收到修改稿时间:2011-04-17

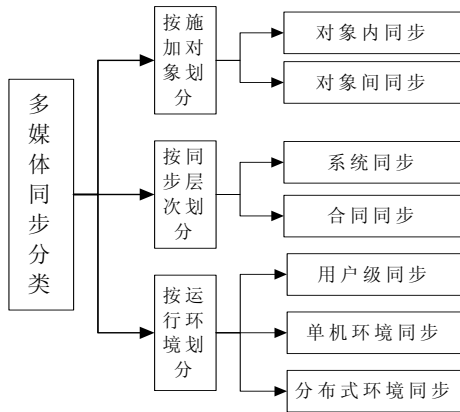


图 1 多媒体同步分类

2.2 音视频同步的影响因素

在互联网环境下，多媒体信息从发送端需要经过一段距离才能到达接收端。在传输过程中，由于受到各种因素的影响，导致在接收端不能正确播放。我们这里列出影响音视频同步的主要因素^[7]，如图 2 所示。

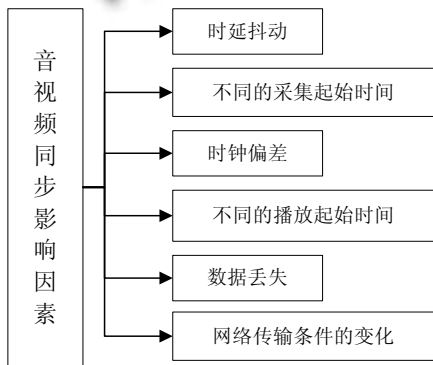


图 2 音视频同步影响因素

3 音视频同步的关键技术

多媒体同步以及音视频同步这些词由来已久，但真正做到同步又能保证音视频质量却不容易。下面介绍几种音视频同步技术及其优缺点，如表 2 所示。

4 音频嵌入视频同步技术

针对当前存在的音视频同步技术的优缺点，音频嵌入视频同步技术显示出了明显的优势，下面介绍一下马晨、陈贺新在《基于 AVS 的嵌入式音视频同步编码器》中提出的基于 H.264/AVS 嵌入式音视频同步技术^[12,13]。

4.1 概述

H.264 是当前主流的视频编码标准，应用前景十

分广泛。基于 H.264/AVS 嵌入式音视频同步技术，利用 DCT 系数奇偶性的音视频同步压缩方法，在进行视频编码时将音频比特流嵌入，对音视频信息进行压缩和传输，如图 3 所示。在解码端提取视频信息时同时提取音频流。此方法对音频质量没有影响，对视频质量影响较小。

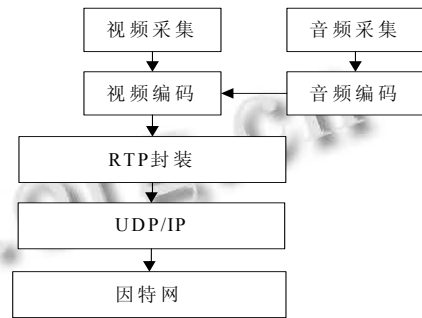


图 3 嵌入式 AVS 音视频同步编码端实时传输模型

4.2 编码端工作原理

1) 嵌入点的选取方法

通过 zig-zag(二维之字型)扫描 8×8 宏块经过 DCT 变换和量化之后的数据，查找最后一个非零点的位置。从 8×8 块的右下角(7, 7)开始到接近左上角的点(2, 0)，即从后向前进行 zig-zag 扫描，直到遇到第一个非零点时停止扫描。如果直到最后一个点时还未停止扫描，则直接将此位置定位到最后一个点;将此位置定义为 C 点。同时将 C 点之前的一点定义为 B 点，这样就定义了两个嵌入点。

2) 做嵌入标记

- a) 如果待嵌入的音频编码比特为 1: 如果 B 点是奇数，则不做改动; 如果 B 点是偶数，则做 B=B+1。
- b) 如果待嵌入的音频编码比特为 0: 如果 B 点是奇数，则做 B=B+1; 如果 B 点是偶数，则不做改动。

$$BQ(u_B, v_B) = BQ(u_B, v_B) + 1 \text{ if } Au_i = 1 \text{ } BQ(u_B, v_B) \% 2 = 0, \\ \{ \\ BQ(u_B, v_B) = BQ(u_B, v_B) \text{ if } Au_i = 0 \text{ } BQ(u_B, v_B) \% 2 = 1$$

3) 算法描述其中 $BQ(u_c, v_c)$ 是 8×8 块中通过 zig-zag 扫描的最后一个非零系数点值; $BQ(u_B, v_B)$ 是 8×8 块中通过 zig-zag 扫描的最后一个非零点前面一个点的系数值; Au_i 代表要嵌入的音频比特值。算法流程图如图 4 所示。

表 2 几种音视频技术的原理及优缺点

音视频同步技术		原理	优点	缺点
音视频媒体内同步技术	基于缓存区数据控制的同步	信宿端改变信宿时钟频率；信源端调整视频流的分辨率。	可以动态适应端到端时延的变化。	改变了媒体流的发送或播放过程。
	基于播放时限的同步	将时延的变化部分通过播放缓冲器进行调整，对抖动进行过滤。	方法简单。	要求端到端时延的上限已知或统计特性可知；不适用于实时通信。
音视频媒体间同步技术	多路复用同步技术 ^[8]	将多个媒体流的数据多路复用到一个数据流或一个报文中	简化媒体流间的同步；不需要附加控制信道和同步化时钟。	浪费带宽资源；不适合媒体流来源于不同结点的情况。
	时间戳同步技术 ^[9]	将媒体数据按时间顺序作时间戳，相同时间戳数据同时表现。	不需要附加信道，不需要同步信息，不改变数据流。	需要全网同步化时钟；读写时间戳开销大。
	同步信道技术 ^[10]	媒体在分离的信道传输，同步信息通过同步信号来单独传输。	能支持复杂的同步关系；可用于直连设备；不需要时钟同步。	同步信息可能丢失或必要同步的媒体数据迟到；开销大。
	同步标记技术 ^[10]	接收方缓存数据直到所有“同步标记”到达时才提交客户。	原理简单；能与现有协议体系良好地兼容。	只支持并行表现；只适合于所有数据来自一个源点的情况。
	基于反馈的同步技术 ^[2]	根据接收端失调检测信息，在发送端或接收端进行同步控制。	适用于网络环境。	具有一定得滞后性；实时性差。
其他技术	时间戳同步信道法 ^[11]	时间戳技术保证流内同步，流间采用同步信道技术。	使流内和流间都有较好的准确性和容错性。	暂未发现
	基于 RTP 协议的同步机制 ^[3]	利用 RTCP 的收、发报文来判断网络的 QoS。	支持异构网络环境；实时性好。	暂未发现
	音频嵌入视频同步技术 ^[12]	根据 DCT 系数特点将音频数据嵌入视频数据中。	开销减小；对音频质量无影响；对视频质量影响也较小。	暂未发现

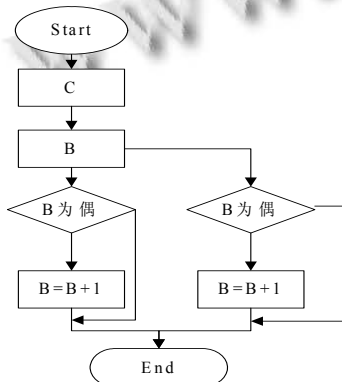


图 4 音频嵌入视频编码端流程图

4.3 音频嵌入视频同步方法的优势

1) 这种同步方法使音频信号隐藏于视频图像中，传输视频同时传输音频，实现了音视频在信道的同步传输，在解码端实现了音频信号的同步到达。

2) 这种同步方法对音频压缩码流没有任何的损失，能够在接收端准确提取音频码流。

3) 这种同步方法对视频的影响：将音频嵌入视频，一定会带来对视频的影响。这个影响可以通过客观和主观两个方面来反映。客观的反应就是信噪比 PSNRY 的下降。主观测试，嵌入之后的视频与原始相

比,很难区分。可以做到人眼观察不到的程度。

4) 关于嵌入开销的分析:嵌入开销的计算:将音频嵌入视频所带来的开销,就是在将音频嵌入视频后生成码流的大小与原始不嵌入时文件的大小之差,再减去嵌入音频的数据量。

嵌入开销的计算如公式所示:

$$Oe = \frac{He - Ov - Ae}{Ov + Ae} \times 100\%$$

经过对多个视频序列的测试,嵌入开销最大是2.0916%,最小只有0.1256。这与通常应用的时间戳技术,所带来的大的系统开销3%相比有了很大的下降。开销基本可以下降一半。

5 结语

音视频同步是多媒体通信中极为重要的内容,本文对与音视频同步相关的多媒体基本概念进行了阐述,并主要介绍了音视频同步的相关技术原理及优缺点。相比而言,时间戳同步技术在时钟同步性好时因其原理简单同步效果好而具有明显优势,而音频嵌入视频同步方法在减小系统开销及嵌入开销方面比时间戳同步技术有明显优势。

虽然音视频同步的研究自从多媒体通信开始出现就已开始,但在如何改进同步效果,减少系统开销,增强实时性方面仍需继续研究。

参考文献

- 1 Escobar J, Deutsch D, Partridge C. Flow Synchronization Protocol. GLOBECOM92. Orlando, Florida: IEEE Computer Society,1992.1381-1387.
- 2 Rangan VP, Ramanathan S, Harrick, et al. Techniques for

Multimedia Synchronization in Network File Systems. *Computer Commun*,1993,16(3):168-176.

- 3 Kuo CC, Chen MS, Chen JC. An adaptive transmission scheme for audio and video synchronization based on real-time transport protocol. 2001 IEEE International Conference on Multimedia and Expo (ICME2001), Tokyo, 22-25 Aug. 2001: 403-406.
- 4 Palacharla S, Karmouch A, Mahmoud SA. Design and Implementation of a Real-time Multimedia Presentation System using RTP. *Computer Software and Applications Conference, (COMPSAC'97)*. 1997: 376-381.
- 5 Chen T. Lip Synchronizatoin Using Speech-Assisted Video Processing. *IEEE Signal Processing Letters*, Apr, 1995,2(4), 57-59.
- 6 钟玉琢,蔡莲红,李树青,等.多媒体计算机技术基础及应用.北京:高等教育出版社,2000.
- 7 董春兵,迟学芬.音视频同步的研究与实现.2007,6:13-15.
- 8 Chen HY, Liu NB, Shiah CW. A Novel Audio/Video Synchronization Model and Its Application in Multimedia Authoring System. *IEEE Trans. on Computer Electronics*, 1995,41(1): 176-177.
- 9 Qazi NU, Woo M, Ghafoor A. A Synchronization and Communication Model for Distributed Multimedia Objects. *ACM Press*.1993:147-155.
- 10 Shepherd D, Salmony M. Extending OSI to Support Synchronization Required by Multimedia Applications. *Computer Communication*, 1990,13(7): 399-406.
- 11 蔡燕.分布式环境下多媒体同步通信系统的研究与设计.武汉:华中科技大学,2001.
- 12 郭尧,陈贺新.基于 AVS 的嵌入式音视频同步传输系统设计.2009,4:45-49.
- 13 马晨.基于 AVS 的嵌入式音视频同步编码器.长春:吉林大学,2009.