

# 基于 PSTN 的远程多媒体监控系统

戴礼荣 李 枫 王仁华

(中国科学技术大学电子工程与信息科学系, 合肥 230027)

**摘 要** 基于 PSTN 网的远程多媒体监控技术,其关键是针对监控应用要求,在保证一定的图象和语音质量的同时,对图象与语音以尽量高的压缩率进行压缩,以及对图象与语音压缩进行实时处理;在所介绍的基于 PSTN 的远程多媒体监控系统中,图象和语音压缩处理是采用 H. 263 和 G. 723.1 协议. 首先介绍了该系统组成结构;然后在简要介绍基于 H. 263 和 G. 723.1 协议的图象与语音压缩算法的基础上,进一步讨论了针对远程多媒体监控的应用要求,基于图象与语音压缩的若干考虑,着重讨论和介绍了该系统的关键组成部分,即多媒体远程监控终端机的设计与实现,提出了基于高速 DSP 技术的图象与语音压缩实时处理的关键技术及所采取的措施,并对系统的主要性能进行了概略的介绍.

**关键词** 多媒体 图象压缩 语音压缩 数字信号处理器

中图分类号: TP911 文献标识码: A 文章编号: 1006-8961(2001)06-0594-05

## PSTN Based Remote Multimedia Surveillance System

DAI Li-rong, LI Feng, WANG Ren-hua

(University of Science & Technology of China, HeFei 230027)

**Abstract** In the application of remote multimedia surveillance based on PSTN(Public Switching Telephone Network), one of the key technologies is the compression of video and speech data stream at a ratio as high as possible, while the quality of the video and speech are ensured for the surveillance application; Other key technologies are such as the real time processing of the compression of video and speech data stream. The PSTN based remote multimedia surveillance system, which is introduced in the article, compresses the video data stream based on the protocol H. 263 and speech data stream based on the protocol G. 723. 1. The article first introduces the architecture of the system. Then after a short introduction to the H. 263 and G. 723. 1 protocols based video and speech compression algorithms, some considerations with the video and speech compression are discussed for the requirements of the remote multimedia surveillance application. Further, the design and realization of the remote surveillance terminal, which is the key part of the system, are introduced with details. Also the key technology and method of the real time processing of speech and video compression based on DSP(Digital Signal Processor) are presented. As a conclusion, the main performances of the system are reported at the end of the article.

**Keywords** Multimedia, Video compression, Speech compression, DSP

## 0 引言

基于 PSTN(公众电话网)的远程监控已在各行各业中得到了广泛应用,由于一般传统的基于 PSTN 的远程监控主要传输一些测量和控制数据,因此对数据传输速率要求不高,但随着多媒体技术

的应用,则产生了对基于 PSTN 多媒体远程监控的要求. 这种多媒体远程监控不仅要通过 PSTN 传输测量和控制数据,以实现测量点的监视和对监控点的控制,而且要传输测控点处的有关实时现场图象和现场语音数据.

尽管 MODEM 数据传输速度已提高到 56kbps,但远远达不到能够实时传输现场图象和现

场语音数据的要求. 因此, 在 PSTN 上实现多媒体远程监控的主要瓶颈问题是数据传输速率问题. 为解决此瓶颈问题, 本文采用符合 H. 324 协议<sup>[1]</sup>的图象和语音压缩处理技术及基于高速 DSP (Digital Signal Processor) 技术, 实现了图象和语音的实时压缩处理.

## 1 远程多媒体监控系统结构

基于 PSTN 远程多媒体监控系统主要由多媒体监控终端机和主控机两大部分组成, 而监控终端机和主控机则通过 MODEM 和 PSTN 来实现物理电路连接. 该系统组成结构如图 1 所示.



图 1 远程多媒体监控的系统组成结构

其中, 多媒体监控终端机的主要功能是: (1) 采集测量点的数据, 并按设定方式将测量数据传送到多媒体监控主机; (2) 接收监控主机的监控命令, 对控制点实施控制; (3) 完成多通道图象和语音信号的切换和实时采集; (4) 进行图象和语音信号的实时压缩处理; (5) 进行多媒体数据的传输控制等.

多媒体主控机的主要功能是: (1) 监控点测量数据处理、分析和控制; (2) 图象和语音信号的实时解码; (3) 对多媒体监控终端机的参数进行设定等处理. 这种多媒体主控机由 PC 机构成, 通过运行基于 Windows 95/98 平台上的监控管理软件, 来完成各种多媒体监控及其管理功能.

多媒体主控机和多媒体监控终端机均可通过主动呼叫来建立通信连接, 如多媒体监控终端机主动呼叫, 则其即按程序设定方式传输图象、语音和测控点监控数据; 如多媒体主控机主动呼叫, 终端机则按主控机控制命令所要求的方式来传输图象、语音和测控点监控数据.

## 2 图象和语音压缩算法及其在多媒体监控应用中的考虑

多媒体监控的数字图象压缩算法采用 H. 263 视频编解码算法<sup>[4]</sup>, 语音压缩算法则采用 G. 723. 1 语音编解码算法<sup>[3]</sup>.

### 2.1 H. 263 视频编解码算法

H. 263 是 ITU-T 提出的作为 H. 324 终端使用的视频编解码建议<sup>[4]</sup>, 它是采用基于运动补偿的 DPCM 混合编码方法, 即首先在运动搜索的基础上进行运动补偿, 然后运用 DCT 变换和“之”字形扫描游程编码, 来得到输出码流. 由于 H. 263 是在 H. 261 协议的基础上, 将运动矢量的搜索增加为半像素点搜索, 同时又增加了无限制运动矢量、基于语法的算术编码、高级预测技术和 PB 帧编码等 4 个高级选项, 从而达到了进一步降低码速率和提高编码质量的目的.

### 2.2 G. 723. 1 语音编解码算法

G. 723. 1 是 ITU-T 提出的作为 H. 324 终端使用的音频信号的压缩编解码算法建议. 它有 5. 27kbps 和 6. 3kbps 两种编解码速率, 且这两种速率的编解码器均有很高的语音质量和较低的时延, 并允许在每帧边界 (30ms) 处采用 5. 27kbps 或 6. 3kbps 两种可变比特率进行传输, 其中编码器大致包括线性预测系数提取和线谱对分析、开环基音值估计、联合滤波器的形成、闭环基音值估计、激励码本搜索等 5 个部分; 而解码器则主要包括 ITU-T G. 723. 1 码流解码、语音信号的 LPC 合成、基音和共振峰后置滤波等 3 个部分.

### 2.3 视频和语音压缩算法应用于多媒体监控的考虑

一般在多媒体监控应用中, 对视频图象质量要求较高, 因为 H. 263 视频编码在图象大面积运动时, 极易产生方块效应, 为解决此矛盾, 视频图象编码采用了 CIF 和 QCIF 两种格式, 其中 CIF 格式按 H. 263 协议的帧内视频编码, 其能够实时编码而不能实时传输, 但能提供高质量的视频图象; 而 QCIF 格式按 H. 263 协议的动图象视频编码, 则既能实时编码, 又能实时传输.

在非实时传输 CIF 格式视频图象时, 需禁止语音传输, 以提高 CIF 格式视频图象的非实时传输速度. 在其他情况下, 则需保证语音的实时性, 并尽量减少语音的延迟.

## 3 多媒体监控终端机结构及视频和语音的实时压缩处理

### 3.1 多媒体监控终端机结构

多媒体监控终端机的硬件总体结构如图 2 所示. 其中, 浮点数字信号处理器 ADSP-21061<sup>[5]</sup>作为多媒

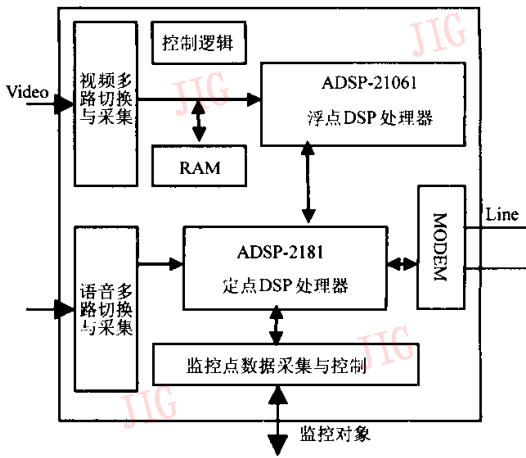


图2 多媒体监控终端机结构框图

体监控终端机的视频图象压缩协处理器,以用于控制视频图象采集,并实现视频图象的实时压缩处理;而定点数字信号处理器 ADSP-2181<sup>[6]</sup>作为多媒体监控终端机的主控处理器,在完成音频采集控制和语音的实时压缩处理的同时,还实施对监控点的数据采集与控制、对 MODEM 的通信控制、对主机交互命令的分析与控制、将图象和语音编码数据按一定格式进行打包处理后,再通过 MODEM 送到远端主机的通信处理、ADSP-21061 数据传输控制,以及通过 ADSP-21061 对图象采集和编码格式等进行控制。

### 3.2 视频图象采集与视频图象的实时压缩处理

#### 3.2.1 视频图象采集的 DMA 控制

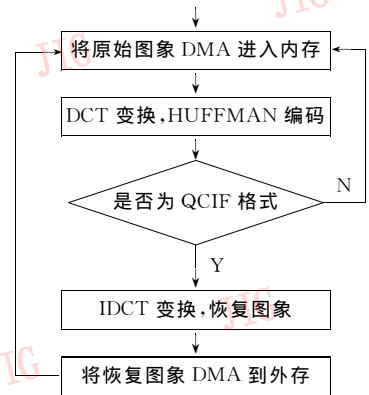
视频图象采集采用 BT819 图象编码器。这里 BT819 作为 ADSP-21061 的外部设备 (External Device),而图 2 中的 RAM 则作为 ADSP-21061 的外部存储器,这是因为 ADSP-21061 的内部 RAM 空间有限,视频图象数据在压缩编码之前需由 ADSP-21061 控制从外部设备 BT819 暂存到外部存储器 RAM 中。另外由于 ADSP-21061 中的 DMA7 具有外部设备到外部存储器的 DMA 控制能力,因此当 BT819 的数据通道采用 API(异步图象接口)的模式 B 时,通过对 BT819 数据接口信号 AEF、AFF、FIELD、RDEN、QCLK,及 ADSP-21061 中的 DMAR 和 DMAG 等信号进行适当的译码和控制,并对 DMA7 通道的参数进行正确的程序设置后,即可用较简单的控制逻辑来实现 ADSP-21061 对视频图象采集的控制,而对于 PAL 制和 NTSC 制 CIF 和 QCIF 格式视频图象数据来说,为保证 BT819 的 FIFO 中的图象数据不丢失,则 DMA7 的外部设备到外部存储器的 DMA 数据传输速度不应低于

8M×16bit/s。

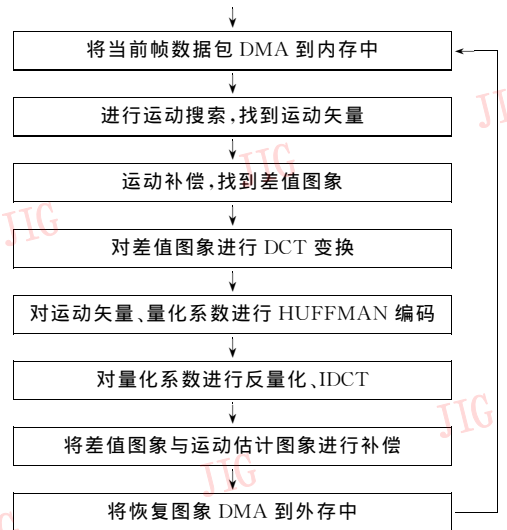
由于 ADSP-21061 DMA 通道的控制是独立的,故 ADSP-21061 对视频图象采集的控制基本上不影响其运算速度。

#### 3.2.2 视频图象实时压缩处理

由 2.3 节的讨论可知,若采集的图象是 CIF 格式,则仅需进行帧内编码;若是 QCIF 格式,不仅需要帧内编码,而且还需帧间编码。对于帧内编码,则不需恢复编码图象;而对于帧间编码,则需恢复编码图象。在 ADSP-21061 上用软件实现的视频图象实时压缩处理,其帧内编码和帧间编码的流程如图 3(a)、(b)所示。



(a) 帧内帧编码流程图



(b) QCIF 的帧间帧编码

图3 视频图象实时压缩处理流程图

在用软件实现视频图象压缩处理算法,即用 H.263 视频编解码协议时,充分考虑和灵活利用 ADSP-21061 很多数字信号处理的优良特性,是视频图象压缩处理算法实时实现的关键。这些优良特性包括多个通道的独立 DMA 控制功能、单周期执

行的多功能指令、片内内置有较大的高速内存、定点运算速度大大高于浮点运算、能进行零开销硬件循环控制等。以下是在考虑到 ADSP-21061 数字信号处理优良特性的基础上,在硬件和软件上所采取的一些主要措施。

(1) 图象数据在存储器中的组织与单周期多功能指令的运用

为节省和有效地利用 ADSP-21061 的内部 RAM 空间,因而将数据量大,且暂时不用的图象数据存放在外部 RAM 中,即由 DMA7 控制,将通过 BT819 采集的 YUV422 格式的原始图象、参考帧图象、恢复帧图象(YUV12 格式)存放在外部 RAM 中。同时在内部 RAM 设置一个块大小为  $16 \times 16 \times 2$  的内存以用于存放从外部 RAM 一次取得的待编码的原始图象数据块,其格式是 YUV422,这时当编码一帧 QCIF ( $176 \times 144$ ) 图象,则需要从外部 RAM 传输  $11 \times 9$  次图象数据。另外,在内部 RAM 设置一个 48 行图象数据大小为  $176 \times 48 \times 1.5$  的内存空间以用于存储帧间帧编码时,运动搜索所需要的参考图象(当编码一帧 QCIF 帧间帧时,需要 3 次数据传输);同时再设置一个块大小为  $16 \times 16 \times 1.5$  的内存空间,用于存取编码后的恢复帧图象(当编码一帧 QCIF ( $176 \times 144$ ) 图象时,需要向外部存放  $11 \times 9$  次恢复图象数据)。

对图象数据组织的主要考虑因素有如下两点:一是在保证 ADSP-21061 高速运行的同时,降低对外部 RAM 的要求,此时若将参加运算的数据存放于 ADSP-21061 的内部高速 RAM 中,则可保证 ADSP-21061 的指令能单周期全速运行,可是如将参加运算的数据存放于外部 RAM 中,则为了使 ADSP-21061 的指令能单周期全速运行,必须要求外部 RAM 的速度达到 ADSP-21061 的要求,否则,需增加等待周期,即降低 DSP 的运行速度;二是使有效使用单周期多功能指令成为可能。由于 ADSP-21061 仅在片内采用 SUPER HARVARD 结构,故只有在数据存放在片内 RAM 中,才能有效使用单周期多功能指令,例如,若两个数据存放在片内,在一条单周期多功能指令中,则能在完成取这两个片内数据的同时,还能完成乘-累加运算等,从而大大提高处理速度。

(2) 外部 RAM 和 ADSP-21061 内部 RAM 之间图象数据的 DMA 传输

若将待处理的图象数据存放在外部 RAM,而

将参加运算的图象数据存放在内部 RAM,则外部 RAM 和内部 RAM 之间图象数据需要较频繁地传输,当需 DSP 的 CPU 参与这种数据的传输时,则必将降低 DSP 的运算处理能力。若此时利用 ADSP-21061 独立的 DMA 控制能力,采用 ADSP-21061 的 DMA6 来实现外部 RAM 和 ADSP-21061 内部 RAM 之间图象数据的传输,并使图象数据传输和运算处理作适当配合,则可做到不影响 ADSP-21061 的运算处理速度。

(3) 数据格式的安排

ADSP-21061 具有较灵活的存储器组织结构,并提供了三种存储器组织方式,即 32bit 的存储器组织,40bit 的存储器组织,16bit 存储器组织。

图象数据一般是 8bit 的非负整型数,对于灰度而言,图象值是  $16 \sim 254$ ;对于色度信号值而言,则是  $16 \sim 240$ ,因而在做运动补偿时,图象数据有可能是负值,这时需要采用 16bit 的短整型数的形式;而在做 DCT 和 IDCT 变换时,为了提高精度,则需采用 32bit 浮点型式的数据。数据格式之所以如此安排,一方面是为了提高了运算速度,另一方面是为了保证运算精度。

### 3.3 语音采集与语音实时压缩处理

语音采集是采用 NS 公司的 A-Law COMBO CODEC 器件,因为它能与 ADSP-2181 的同步串口进行无缝连接,ADSP-2181 的串口初始化为位同步时钟 2.048MHz,帧同步 8kHz,A-Law 压扩。ADSP-2181 是 16bit 定点 DSP,指令周期为 30ns,运算速度为 33MIPS(百万条指令每秒),而且所有指令单周期执行,片内有  $24\text{bit} \times 16\text{k}$  的程序 RAM (PRAM)和  $16\text{bit} \times 16\text{k}$  的数据 RAM (DRAM)。其多功能指令在单周期内能够执行一次乘-累加运算和存取 PRAM 和 DRAM 的数据各一个。在 ADSP-2181 上实现 G.723.1 语音实时压缩处理所采用的双缓冲数据流程图 4 所示。

考虑到 ADSP-2181 的硬件资源和其指令特点,在 ADSP-2181 上用软件来实现实时语音的压缩处理时采取了一些措施,如,在数据组织上,尽量采用符合其多功能指令要求的数据组织方式,即在数据变量的存储空间分配上,为节省空间,对模块内的临时变量采用覆盖的方法,以避免分配新的存储空间;在数据格式处理上,对精度要求比较高的运算中间结果,采用高比特数据格式(如 32bit),而运算的最

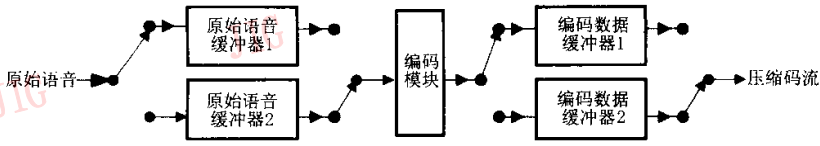


图4 语音压缩双缓冲数据流程

后结果仍采用 16bit 格式表示。

### 3.4 ADSP-2181 与 ADSP-21061 的通信

ADSP-21061 与 ADSP-2181 的通信主要是编码后的图象数据和控制命令的传送。虽然实现它们之间的通信有多种可选方案,但考虑到它们之间的数据传输量不大,因此最简洁的实现方式是通过它们的同步串口来实现。ADSP-21061 与 ADSP-2181 均具有硬件支持的两个可编程同步串口,因此在适当编程的前提下,它们的同步串口可实现无缝连接(如图 5 所示)。

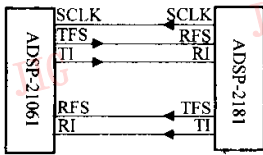


图5 DSP 间的通信

事实上,在 ADSP-2181 和 ADSP-21061 之间所进行的图象数据和命令的传输是一种异步传输,即当 ADSP-2181 需要向 ADSP-21061 取得数据或者状态时,就向 ADSP-21061 发出一个命令,而 ADSP-21061 接到命令后就向 ADSP-2181 报告状态或者发送数据;也就是说,两者是通过握手来取得数据的通信。为避免耗费 DSP 的 CPU 时间去过滤由串口同步工作所产生的垃圾数据,可将软件控制串口的发送和接收转换为伪非同步通信方式,以实现所要求的图象数据和命令的异步传输。

## 4 结 语

本文讨论的基于 ADSP-21061 和 ADSP-2181 的多媒体远程监控系统,其图象压缩实时处理能力达到: CIF 格式帧内编码为 7 帧/s,码流大小为 48~64kbps; QCIF 格式动图象编码为 10 帧/s,码流大小为 22kbps。语音压缩实时处理占用 ADSP-2181 约 67% CPU 的时间,码流为可选的 5.27kbps 或 6.3kbps。

由于 DSP 在性能提高的同时,价格也在下降,而且事实上, DSP 的价格已逼近微控制处理器或相应

的 ASIC 电路的价格,如 ADSP-21061 的价格在 10 千只批量时,约为 10 美元,所以,基于 DSP 的系统同样会取得很高的性能价格比。如本文所介绍的远程多媒体监控系统。另外,由于 DSP 的可编性,基于 DSP 的系统还容易实现产品的升级和功能的改善。

## 参 考 文 献

- 1 ITU-T Recommendation H.324: terminal for low bitrate multimedia communication. 1996.
- 2 王仁华,郭武,李锦宇等. PSTN 网上可视多媒体通信终端. 通信学报,1998,20(5):41~49.
- 3 ITU-T Recommendation G.723. Dual rate speech coder for multimedia communications transmitting at 5.3&6.3kbit/s. Oct,1995.
- 4 ITU-T Recommendation H.263: Video coding for low bitrate communication. 1995.
- 5 Analog Devices. Inc. ADSP2106X SHARC User's Manual. New Jersey:Prentice-Hall,1994.
- 6 Analog Devices. Inc. ADSP2100 Family User's Manual. 2nd ed. New Jersey:Prentice-Hall,1996.

戴礼荣 1962 年生,1997 年获中国科学技术大学博士学位。主要从事语音信号处理、多媒体通信、语音分离与语音识别的研究以及高速 DSP 硬件、软件系统开发与应用。

李 枫 中国科技大学电子工程与信息科学系硕士生。研究方向主要有多媒体通信、数字视频处理及实时信号处理。

王仁华 中国科学技术大学电子工程与信息科学系教授,博士生导师。中国通信学会会员、理事。主要从事数字信号处理、语音通信、多媒体通信等方面的研究。