



**摘要:** 本文讨论了一种在PSTN网上实现基于PC的实时视频通信的方法, 它将对在窄带通信网上实现可视电话、移动可视通信、远程监控等业务有重要的意义。

**关键词:** 视频通信 H.263 线程 同步

### 引言

随着通信技术的不断发展, 特别是多媒体通信的发展, 如何在现有的网络上很好地进行视频通信已成为人们日益关注的焦点之一。为了在PSTN上实现多媒体通信, ITU-T于1996年提出了H.324系列建议(多媒体通信终端建议), 其中H.263是其所采用的视频压缩标准, 它使得在现有窄带通信网络上实现实时视频通信成为可能。

本文在简单介绍基于PC的H.263视频编码实际系统的基础上, 着重讨论了软件编码优化方法、多线程设置和Modem通信等问题。采用本文所述的方法已通过Modem在PSTN网上实现了实时的视频通信, 获得了较好的图像质量。

### 系统框图

在PSTN上实现实时视频通信的系统框图如图1所示:

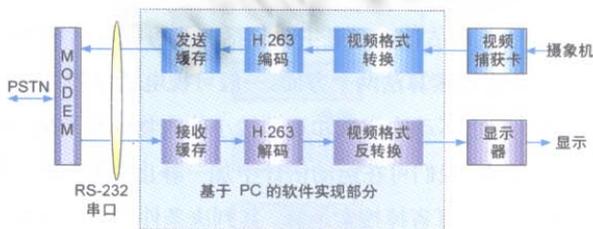


图 1 实时视频通信的系统框图

如图1所示, 首先视频捕获卡将摄像机输出的模拟或数字信号送到视频格式转换部分, 变为H.263编码所需的YUV411格式的数据后放入相关的缓存中。然后, H.263

编码器从缓存中取出数据进行编码, 并将编码后的数据流放入发送缓存中。然后, 发送缓存中的图像数据通过串口后由Modem送到PSTN网络中传输给通信对方。同时, 系统还从PSTN网络中接收通信对方的数据, 并放入接收缓存中。随后H.263解码器从接收缓存中取出数据解码, 并送到计算机显示器去显示。这样通信双方就在PSTN上实现了视频通信。

### 软件编解码

由系统实现框图可知, 本系统的主要部分为软件视频编解码。我们是利用Windows操作系统中由Win32提供的抢先式多任务并行机制来实现该部分的。同时, 为了满足视频通信对实时性的要求, 我们采用多线程的方式来构造程序的主体框架。

系统用软件实现的框图如图2所示:

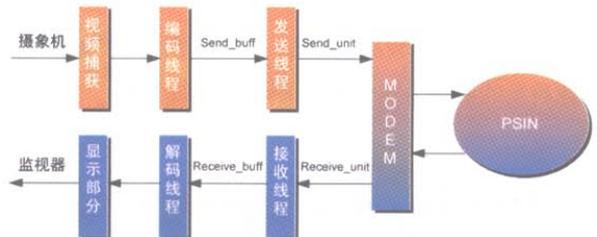


图 2 实时视频通信系统的实现框图

根据框图, 在程序设计时建立四个线程分别完成编码、发送、接收、解码及显示等功能。以下为各个部分的具体功能:

**视频捕获:** 完成视频图像的采集, 将捕获到的一帧数据以 Y、U、V 三数组方式存放在固定的内存中。

**编码线程:** 对视频捕获采集到的图像数据进行优化的 H.263 编码, 并将编码结果送到发送缓冲区 Send\_buff。

**发送线程:** 从发送缓冲区 Send\_buff 中每次取出一定量的数据放入发送单元 Send\_unit 中发送给 Modem, 这样可以保证每次向 Modem 写入一定量的数据。

**接收线程:** 从 Modem 每次读取与发送单元大小相等的的数据放入 Receive\_unit 中, 再将 Receive\_unit 中的数据存入接收缓冲区 Receive\_buff 中。

**解码线程:** 将 Receive\_buff 中的数据取出进行解码。

**显示:** 将解码结果输出到监视器显示。

### 几项关键技术

为了增强视频通信的实时性以及获得较好的图像质量, 本文在以下几个方面做了改进:

#### 1. H.263 编码的优化

由于视频编解码速度的快慢会直接影响到整个视频通信的性能, 所以必须对 H.263 编解码算法进行有效的优化, 以达到实时编解码的目的。在实现过程中具体采用了三种不同级别的优化: 框架级优化、局部算法优化和代码优化。

(1) 框架级优化: 框架级优化主要是对程序总体框架和算法的全局控制的优化。由于编码过程中帧间编码函数所占的 CPU 时间最多, 所以从全局的角度对其优化是很有必要的, 可以避免大量不必要的运算。

例如, 在帧间编码中, 如果预测误差块中全零, 则称此预测误差块为“全零块”; 如果预测误差块经过 DCT 和量化后, 量化系数块全零, 则称此预测误差块为“准全零块”。对于运动较小的图像序列, 出现“全零块”或“准全零块”的概率比较高, 为此, 可以利用这一现象在保证图像质量的前提下减少帧间编码所需的时间。

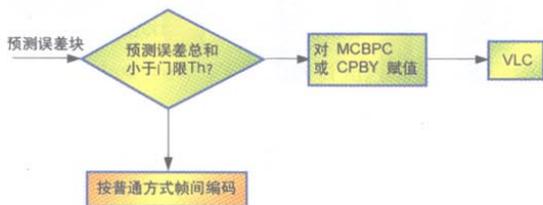


图 3 改进后的帧间编码框图

在帧间编码中主要是 DCT 和量化比较费时, 若在 DCT 和量化之前就知道当前预测误差块是“全零块”或“准全零块”, 则不必进行 DCT、量化及图像重建中的反量化和 IDCT, 只需对 H.263 码流中相关的码字如 MCBPC(宏块类型与编码块色差类型)或 CPBY(亮度编码类型)进行相关的比特赋值即可。所以改进后的帧间编码算法, 在 DCT 前加一“全零块”或“准全零块”的判断, 若条件不满足, 则进行普通的帧间编码, 反之则只进行相关的比特赋值和简单的宏块头编码。这样就减少了帧间编码的运算量从而提高了编码速度。改进后的帧间编码框图如图 3 所示:

(2) 算法级优化: 局部算法的优化主要是对帧内、帧间编码中涉及到的耗时较多的一些算法如 DCT、运动估计等算法进行优化。

对于 DCT/IDCT 方面, 可以通过利用输入视频序列的特点、DCT/IDCT 行列变换的规律及人眼的视觉特性来减少 DCT/IDCT 的运算量。对于运动较小的图像序列, 其帧间编码中的预测误差块的同一行(或列)中各个元素“全等”的概率比较大。同时又由于一维 DCT 变换具有以下规律: 若输入的一维数组的元素全等, 则经变换后 AC 系数全为零, DC 系数为输入的各元素的累加和。因此利用这些特点和规律, 可在解码端对变换系数块进行行(或列)IDCT 之前, 首先判断 AC 系数是否为全零, 若是, 则不用运算, 只需将 DC 系数除以元素个数再赋给输出的各个元素即可。同理, 在编码端对预测误差块进行行(或列)DCT 时, 可首先对该行中元素是否全等进行判决, 同时考虑到人眼视觉特性, 我们还可将两个元素 A、B 是否相等的条件放宽为:  $|A-B| < Th$ , 其中 Th 是一个可选择的门限值。这样可增加出现 AC 系数为零的概率, 从而减少了运算量。

对于运动估计, 改进主要包括利用可视电话序列的规律和改进搜索算法两个方面。一般可视电话序列具有如下规律: 静止不动的背景和运动较小的人物的头像像。利用这一规律, 我们可在运动估计中加一静止块的判断, 若为静止块, 则可省掉搜索运算。其判决条件为: 将预测误差块元素的总和和同一门限比较, 若小于该门限, 则判为静止, 否则判为运动。另一方面, 在研究和实现运动估计的快速搜索算法后, 我们发现这些算法都是通过如下方式来减少搜索所需的计算量的:

- 利用可视电话运动基本上是中心分布的规律;
- 在搜索过程中采用半途停止技术来减少搜索步骤;

· 去除不必要的搜索点。

我们的快速运动估计算法 SES-BBGDS(简单有效的基于块的梯度下降搜索算法)综合利用了上述几点,可将搜索点数减少 3~5 个。

### (3) 代码级优化

代码级的优化主要是采用高效的语言来实现一些程序代码如汇编等。它包括:利用编码技巧,如将一些短循环语句解开、定义恰当的函数参数以减少参数传递所花的时间、避免重复的指针地址运算等技巧,以减少运算量和程序运行所花的时间;将部分程序代码采用 MMX 汇编指令来实现,从而提高程序运算的速度。

## 2. Modem 通信

在发送端 Modem 将编码后的码流转换成能在 PSTN 上传输的模拟信号,在接收端 Modem 将收到的模拟信号转换成比特流送给解码器。在 Windows 操作系统中提供了通信 API 和电话 API,利用这些 API 函数,可以比较方便的实现通过 Modem 连接到电话线的计算机之间的通信。这将大大减小编程的复杂度。实现 Modem 通信的流程图可由图 4 所示。

创建呼叫后,每次呼叫状态改变时,Windows 都会发送 LINE\_CALLSTATE 消息通知控制应用程序。一旦呼叫进入 CONNECTED 状态,则表明连接已经建立,应用程序就可以通过电话线传送图像数据了。

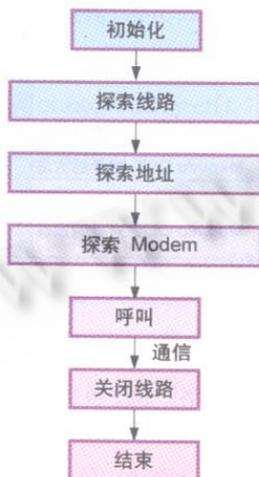


图 4 Modem 通信流程图

## 3. 线程同步和优先级的设置

实现程序中共有四个线程:编码线程、发送线程、接收线程、解码线程,它们之间必须处理好同步问题,否则

程序就会出错。我们采用临界区和事件的方法来实现线程之间的同步。例如,发送线程和编码线程之间可以通过临界区对象来保证对编码后数据的读写不会同时进行,防止数据出错。同时通过事件对象来同步线程,以防止发送线程从发送缓冲区(Send\_buff)中取不到数据,以及编码线程造成发送缓冲区内的数据溢出。同理,在接收线程和解码线程之间我们也采用了临界区和对象的方法来实现同步。

由于 Modem 传输速率的限制,在很大程度上影响了视频通信的实时性,为此我们根据各个线程所占 CPU 的时间不同来设定优先级。对于占用 CPU 时间较长的线程,将它的优先级设为较低;反之,对于占用 CPU 时间较少的线程,将它的优先级设为较高。这样就减少了线程相互之间的等待时间,有效地提高了视频通信系统的通信速度。

## 总结

本文以 PC 为平台,通过调制解调器(33.6bps)在 PSTN 上实现了实时的视频通信。所传图像质量及通信效果较好。在 33.6kb/s 的线路上,对于 CIF 图像格式,可以达到 3~4 帧/秒的解码帧频;对于 QCIF 图像格式,可以达到 10~12 帧/秒的解码帧频。同时,该视频通信系统的实现将可推广应用到可视电话、移动可视通信、远程教育、ISDN 上的图像通信、远程监控、基于 IP 的视频通信等多个领域。■

## 参考文献

- 1 TU-T Recommendation H.263(1996):Video coding for low bit rate communication
- 2 [美] Charles A.Mirho&Andre Terrisse 编著 《Windows 95 通信编程》 清华大学出版社
- 3 胡栋 朱秀昌 编著 《图像通信技术及应用》 人民邮电出版社
- 4 Microsoft 公司编著 《Win32 程序员参考大全(2)》 清华大学出版社