

H. 264 视频编码器的研究与分析

郝瑞林

北京邮电大学网络技术研究院 北京 (100876)

E-mail: ruilynn@gmail.com

摘要: 随着市场对多媒体信息传输需求的增加, 多媒体信息传输如何适应不同信道传输特性的问题日益显现出来。H. 264 作为新一代视频编解码标准, 它压缩性能优异, 图像质量高, 可满足各种不同速率、不同场合的视频应用。正因为功能的多样性, 使其编码方法异常复杂, 因此本文分析了 H. 264 视频编码器的原理, 对当今流行的 H. 264 编码器进行了性能比较, 以 X264 开源编码器为例进行了源码级的分析, 详细介绍了 X264 中的函数流程与算法, 并根据具体应用进行了不同参数设置的性能分析。本文的分析能使用户对 H. 264 视频编码原理有更深入的认识, 并为自己的应用设计独特的视频编码器, 提高视频编码性能。

关键词: H.264; X264; 视频编码

中图分类号: TP391

1. 引言

视频编解码技术是流媒体技术的关键技术之一, 从 H.261 视频编码建议, 到 H.262/3、MPEG-1/2/4 等都有一个共同的不断追求的目标, 即在尽可能低的码率(或存储容量)下获得尽可能好的图像质量。而且随着市场对图像传输需求的增加, 如何适应不同信道传输特性的问题也日益显现出来。于是 IEO/IEC 和 ITU-T 两大国际标准化组织联手制定了视频新标准 H.264^[1]来解决这些问题。

H.264/AVC 引入了面向 IP 包的编码机制, 有利于网络中的分组传输, 支持网络中视频的流媒体传输。H.264 具有较强的抗误码特性^[2], 可适应丢包率高、干扰严重的无线信道中的视频传输。H.264 支持不同网络资源下的分级编码传输, 从而获得平稳的图像质量。H.264 能适应于不同网络中的视频传输, 网络亲和性好。

正因为功能的多样性, 使其编码方法异常复杂, 因此本文分析了 H. 264 视频编码器的原理, 对当今流行的 H. 264 编码器进行了比较, 以 X264 开源编码器为例进行了源码级的分析, 详细介绍了 X264 中的主要数据结构和函数调用关系, 并根据具体应用进行了不同参数设置的性能分析。本文的分析能使用户对 H. 264 视频编码原理有更深入的认识, 并为自己的应用设计独特的视频编码器, 提高视频编码性能。

2. H.264 编码器原理

2.1 H. 264 编码器的功能结构

如图 2-1 所示, H.264 标准压缩系统由视频编码层 (Video Coding Layer, VCL) 和网络提取层 (Network Abstraction Layer, NAL) 两部分组成^[3]。VCL 层主要包括帧内预测, 帧间预测、变换量化、熵编码等压缩单元。NAL 层则用于为 VCL 层提供一个与网络无关的统一接口, 它负责对视频数据进行封装打包后使其在网络中传送, 它采用统一的数据格式, 包括单个字节的包头信息、多个字节的视频数据与组帧、逻辑信道信令、定时信息、序列结束信号等。包头中包含存储标志和类型标志。存储标志用于指示当前数据不属于被参考的帧。类型标志用于指示图像数据的类型。VCL 可以传输按当前的网络情况调整的编码参数。



图 2-1 H.264 功能结构示意图

2.2 H.264 编码器工作流程

如图 2-2 所示，编码器采用的是变换和预测的混合编码法。输入的帧或场 F_n 以宏块为单位被编码器处理^[4]。首先，按帧内或帧间预测编码的方法进行处理。若采用帧内预测编码，其预测值 $PRED$ 是由当前片中当前块的左侧和上侧已编码的像素点推出的，而如果采用帧间预测编码，其预测值 $PRED$ 是由已编码的视频帧/场通过运动补偿得到的。预测值 $PRED$ 和当前块相减后，产生一个残差块 D_n ，经变换，量化后产生的变化系数 X 再经熵编码，与解码需要信息一起组成一个压缩码流，经 NAL 供传输和存储用。为进一步提供预测用的参考图像，编码器还需要有重建图像的功能。即将残差图像经反量化，反变换后得到 D_n' 与预测值 $PRED$ 相加，得到 uF_n' ，经过滤波后得到 F_n' 即为重建图像，可用于参考图像。

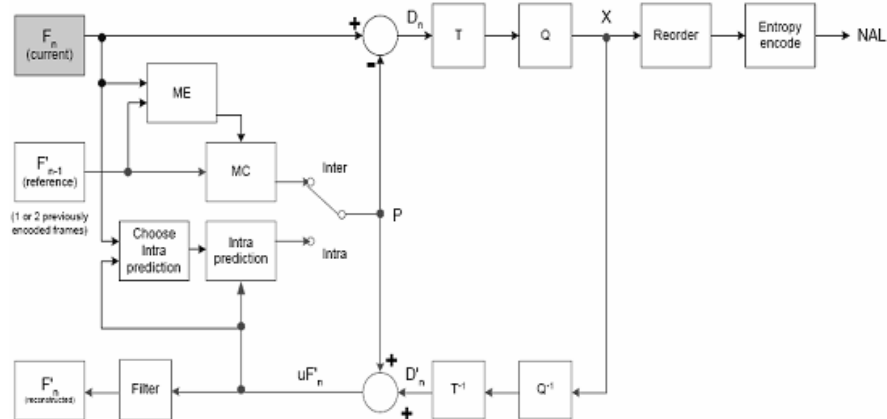


图 2-2 H.264 编码器结构图

3. X264 开源编码器分析

3.1 三大开源编码器

目前 H.264 的开源编码器主要有三类：JM，T264 和 X264。JM 作为 H.264 的官方测试源码，由德国 hhi 研究所负责开发。它实现了 264 所有的特性，但其程序结构冗长，只考虑引入各种新特性以提高编码性能，忽视了编码复杂度，其编码复杂度极高，不宜实用。T264 是中国视频编码自由组织联合开发的 H.264 编解码器，编码器能编码输出标准的 H.264 码流，但解码器只能解 T264 编码器生成的码流，目前已经停止更新。X264 是网上自由组织联合开发的兼容 264 标准码流的编码器，它注重实用，和 JM 相比，在不明显降低编码性能的前提下，努力降低编码的计算复杂度，编码效率很高。

3.2 X264 流程

因为 X264 编码器的高效率和高性能,应用之中较多使用,下面对 X264 进行源码分析,进一步分析 H.264 的编码器原理。

在 X264 编码器的调度中最重要的是 Encode 函数,是对整个视频序列的 H.264 编码。在函数中,首先进行了函数的初始化操作,然后判断是否已经编码完毕,如果没有则继续读取一帧数据到缓冲区,然后进行帧解码;如果编码完毕则退出编码流程。详细函数关系如图 3-1 所示:

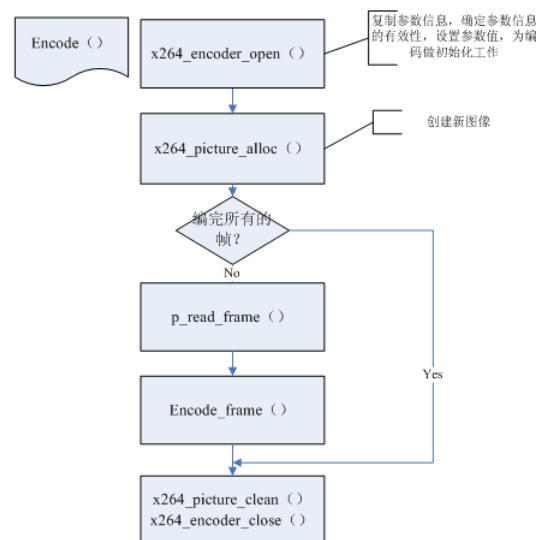


图 3-1 Endcode 函数流程图

其中, `x264_encoder_open()` 这个函数是对不正确的参数进行修改,并对各结构体参数和 cabac 编码、预测等需要的参数进行初始化。`x264_picture_alloc()` 这个函数分配能容纳一帧 `sizeof(x264_picture_t)` 字节数的空间,然后进行初始化。`p_read_frame()` 这个函数就是一次读入一帧到刚分配的空间,这里的数据都是原始的视频图像数据。`Encode_frame()` 这个函数对视频序列其中一帧进行 264 编码。`x264_picture_clean()` 和 `x264_encoder_close()` 两个函数主要是编码后的处理工作,如将帧数据全置零等。

在 `Encode_frame()` 函数中开始上文提到的 VCL 编码和 NAL 编码,其中 `x264_encoder_encode()` 函数为 VCL 层编码,其详细流程如图 3-2 所示。其中, `Setup new frame from picture` 主要是将图片的原始数据赋值给一个未使用的帧,用于编码。`Get frame to be encoded` 主要是帧管理的操作,从编码帧的缓存中取出一帧来对他进行编码。`Setup frame context` 主要是对即将编码帧进行帧类型的预设。`Init`, `Write the bitstream` 主要是对参考列表的初始化,片头的初始化,以及对将编码后数据写入比特流进行传输。`Update encoder state` 这部分是一帧编码后的编码器的更新处理部分,主要有参考帧的管理,去块滤波,像素内插等工作。`Compute/Print statistics` 这部分并不属于编码的工作,只是对其中编码性能的统计计算和显示工作。

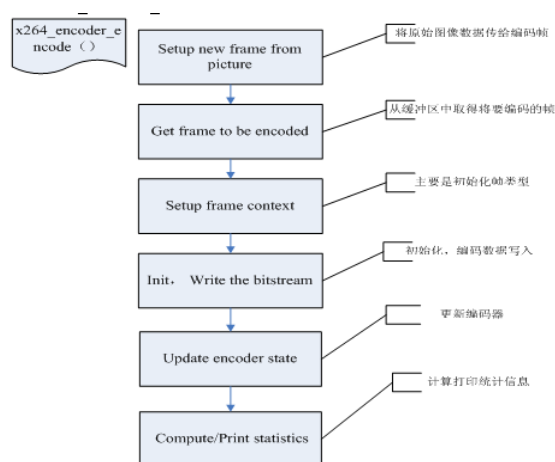


图 3-2 x264_encoder_encode 函数流程图

x264_slice_write()函数是编码中最重要的函数, 以上所介绍的预测编码等都在这个函数中实现的。Init()函数主要是初始化的一些操作。x264_macroblock_cache_load()函数主要是把当前宏块的 up 宏块和 left 宏块的预测模式, 非零系数值等数据加载进来, 放到一个数组里面, 供当前模块参考使用。x264_macroblock_analyse()函数主要是模式选择的问题, 通过对 SAD 值或者其他 COST 值的分析, 确定当前宏块的编码类型。以 I 帧模块为例, 我们可以将它分割成 16 个 4*4 的块, 如果这 16 个块的 sad 加起来小于按 16*16 的方式计算出来的 sad 值, 我们就将这个 16*16 的块分成 16 个 4*4 的块进行编码, 否则采用 16*16 的方式编码。x264_macroblock_encode()函数即是依据上面所确定的编码模式对当前宏块进行 264 编码。CABAC/CAVLC 部分为熵编码部分。

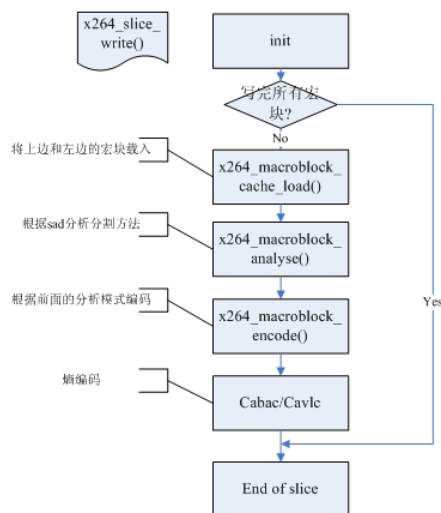


图 3-3 x264_slice_write 函数流程图

3.3 预测编码算法

帧内预测编码的预测值 P 是在一定的预测模式下, 通过邻近已编码的像素值推出的。H.264 中的帧内预测分为 4x4 子块和 16x16 子块以及对 8x8 子块的预测模式。在 x264 代码中, 帧内预测模式的算法主要在 predict.c 中。其中亮度像素有三种方式: 16x16 宏块预测模式, 8x8 子块预测模式, 4x4 子块预测模式; 色度像素有只有 8x8 色度预测模式。

帧间预测编码主要包括运动估计, 运动补偿等, 其中运动估计尤为重要。在 x264 中帧间运动估计有三种算法可供选择 X264_ME_DIA, X264_ME_HEX, X264_ME_ESA。

X264_ME_ESA: :全搜索法,也称为穷尽搜索法,是对搜索范围内所有可能的候选位置计算 SAD(i,j)值,从中找出最小 SAD(绝对差值和),其对应偏移量即为所求运动矢量。此算法虽计算量大,但最简单、可靠,找到的必为全局最优解。

X264_ME_DIA: :菱形搜索^[5],搜索模板的形状和大小不但影响整个算法的运行速度,而且也影响它的性能。块匹配的误差实际上是在搜索范围内建立了误差表面函数,全局最小点即对应着最佳运动矢量。基于这两点事实,菱形算法采用了两种搜索模板,分别是有 9 个检测点的大模板 LDSP(Large Diamond Search Pattern)和有 5 个检测点的小模板 SDSP(SmallDiamond Search Pattern),搜索时先用大模板计算,当最小块误差 MBD 点出现在中心点处时,将大模板 LDSP 换为 SDSP,再进行匹配计算,这时 5 个点中的 MBD 即为最优匹配点。

X264_ME_HEX: :六边形搜索,与菱形搜索相似,只是搜索形状为六边形。

4. 验证分析

针对视频监控这一特殊应用来说,通常都是静止的图像,所以使用相对高速而图像质量下降并不是很多的算法是比较好的。

根据以上的分析,对编码器做如下修改:

1. 帧内预测取消了亮度 18x8 帧内预测模式;
2. 帧间预测运动估计算法使用 X264_ME_DIA; 运动估计的搜索范围由 16 改为 8; 亚像素搜索算法由每一次模式分析时都做一次 1/2 像素搜索,二次 1/4 像素搜索改成只在模式选择之后做一次 1/2 像素搜索,一次 1/4 像素搜索; 模式选择取消了 16x8,8x16,8x4,4x8,4x4 子模块;
3. 变换量化部分取消了 8x8 变换。

测试环境: AMD Athlon XP 1800+ 512M DDR WindowsXP Professional

测试序列: Akiyo, Mobile, Flower

可以得到修改前后的对比:

表 1 psnr 对比

	默认			修改后		
	Y	U	V	Y	U	V
Akiyo	40.21	43.03	43.78	40.03	43.02	43.76
Mobile	34.71	35.68	35.49	34.62	35.67	35.48
Flower	36.23	37.66	38.09	36.13	37.65	38.07

表 2 kbps 和 fps 对比

	默认		修改后	
	Kbps	Fps	Kbps	Fps
Akiyo	32.11	54.08	35.56	63.57
Mobile	485.14	36.50	508.33	63.16
Flower	1694.15	14.11	1802.02	21.11

表 3 cpu 占有率及内存占用

	默认		修改后	
	Cpu 占有率	内存占用	Cpu 占有率	内存占用
Akiyo	35—45%	5235	20—30%	5120
Mobile	55—65%	5240	40—50%	5130
Flower	75—88%	6760	65—75%	6616

5. 结束语

从以上的分析可以看出,经过特殊设定的编码器的编码器复杂度明显下降了,在测试序列 mobile 中编码速度差不多提高了 70%,cpu 占有率也明显的降低了 10—15%,内存的占用也有所降低;而 psnr 的下降都在 0.2db 以下,对人眼来说是基本上没有区别,码率的提升也在 10% 以下。

在特定的视频应用中,根据对编码器的了解,设置正确的参数,往往能极大的提高编码器的性能。

参考文献

- [1] Thomas Wiegand,Gary Sullivan. Overview of the H.264/AVC Video Coding Standard[C]. .IEEE transactions on circuits and systems for video technology,2003-07 ,Vol.13,No.7:560-575
- [2] Iain Richardson. H.264 And MPEG-4 video compression[M]. England: John Wiley & Sons Ltd, 2004
- [3] Keng-Pang Lim ,Gary Sullivan and Thomas Wiegand. Text Description of Joint Model Reference Encoding Methods and Decoding Concealment Methods[R]. Busan,Korea: Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, 2005-11-09
- [4] 毕厚杰. 《新一代视频压缩编码标准——H.264/AVC》[M], 北京: 人民邮电出版社, 2005
- [5] 沈兰荪, 卓力等. 《视频编码与低速率视频传输》[M], 北京: 电子工业出版社, 2001。

The Research and Analysis of H.264 Video Encoder

Hao Ruilin

Beijing University of Posts & Telecommunications, Beijing (100876)

Abstract

As the market demand for multimedia information transmission increased, it's increasingly apparent how to transmit multimedia messages with adapting to different channel transmission characteristics.H.264,as new video encoding standard, which has high compression performance and image quality, can meet the need of the video applications with different rates on different occasions.It is very complex because of the diversity of functions.In this paper,it analyzes the principle of H.264 video encoder, compares the performance of popular H.264 encoder,analyzes the sources of X264 encoder,describes the procedure and algorithm of X264 and analyzes the performance through the specific application with different parameters.In this paper, the analysis will make users more understanding to H.264 video coding principle and improve the performance of video encoder with their own video encoder in their specific application.

Keywords: H.264, X264, video encoding