

流媒体技术笔记：视频编码相关

一、视频编码原理

视频图像数据有极强的相关性，也就是说有大量的冗余信息。其中冗余信息可分为空域冗余信息和时域冗余信息。压缩技术就是将数据中的冗余信息去掉（去除数据之间的相关性），压缩技术包含帧内图像数据压缩技术、帧间图像数据压缩技术和熵编码压缩技术。

去时域冗余信息

使用帧间编码技术可去除时域冗余信息，它包括以下三部分：

—— 运动补偿：

运动补偿是通过先前的局部图像来预测、补偿当前的局部图像，它是减少帧序列冗余信息的有效方法。

—— 运动表示

不同区域的图像需要使用不同的运动矢量来描述运动信息。运动矢量通过熵编码进行压缩。

—— 运动估计

运动估计是从视频序列中抽取运动信息的一整套技术。

注：通用的压缩标准都使用基于块的运动估计和运动补偿。

去空域冗余信息

主要使用帧内编码技术和熵编码技术：

—— 变换编码

帧内图像和预测差分信号都有很高的空域冗余信息。变换编码将空域信号变换到另一正交矢量空间，使其相关性下降，数据冗余度减小。

—— 量化编码

经过变换编码后，产生一批变换系数，对这些系数进行量化，使编码器的输出达到一定的位率。这一过程导致精度的降低。

—— 熵编码

熵编码是无损编码。它对变换、量化后得到的系数和运动信息，进行进一步的压缩。

二、视频编码分类

1、MPEG 系列

（由ISO/国际标准化组织机构/下属的MPEG[运动图像专家组/开发）视频编码方面主要是MPEG1（VCD使用）、MPEG2（DVD使用）、MPEG4（DVD RIP使用的都是它的变种，如：divx，xvid等）、MPEG4 AVC。

1.1、MPEG-1 编码

MPEG-1 是 MPEG 组织制定的第一个视频和音频有损压缩标准。视频压缩算法于 1990 年定义完成。1992 年底，MPEG-1 正式被批准成为国际标准。MPEG-1 是为 CD 光碟介质定制的视频和音频压缩格式。一张 70 分钟的 CD 光碟传输速率大约在 1.4Mbps。而 MPEG-1 采用了块方式的运动补偿、离散余弦变换（DCT）、量化等技术，并为 1.2Mbps 传输速率进行了优化。MPEG-1 随后被 Video CD 采用作为核心技术。

特点：

- 随机访问
- 灵活的帧率
- 可变的图像尺寸
- 定义了I-帧、P-帧和B-帧
- 运动补偿可跨越多个帧

- 半像素精度的运动向量
- 量化矩阵
- GOF结构
- slice结构

缺点（视频监控中）：

- 压缩比还不够大：在多路监控情况下，录像所要求的磁盘空间过大。
- 图像清晰度还不够高：由于MPEG1最大清晰度仅为352 * 288，考虑到容量、模拟数字量化损失等其它因素，回放清晰度不高。
- 不适合网络传输：对传输图像的带宽有一定的要求，不适合网络传输，尤其是在常用的低带宽网络上无法实现远程多路视频传送。
- 灵活性差：MPEG1的录像帧数固定为每秒25帧，不能丢帧录像。

相关 RFC（RTP 包）：

- a) rfc2038 : RTP Payload Format for MPEG1/MPEG2 Video
<http://tools.ietf.org/rfc/rfc2038.txt>
- b) rfc2250 : RTP Payload Format for MPEG1/MPEG2 Video
<http://www.ietf.org/rfc/rfc2250.txt>

1.2、MPEG-2 编码

MPEG-2 制定于 1994 年，设计目标是高级工业标准的图象质量以及更高的传输率。MPEG-2 所能提供的传输率在 3-10Mbps/sec 间，其在 NTSC 制式下的分辨率可达 720X486，MPEG-2 也可提供并能够提供广播级的视像和 CD 级的音质。MPEG-2 的音频编码可提供左右中及两个环绕声道,以及一个加重低音声道，剖析 MPEG-2 视频编码器和多达 7 个伴音声道（DVD 可有 8 种语言配音的原因）。由于 MPEG-2 在设计时的巧妙处理，使得大多数 MPEG-2 解码器也可播放 MPEG-1 格式的数据，如 VCD。

同时，由于 MPEG-2 的出色性能表现，已能适用于 HDTV，使得原打算为 HDTV 设计的 MPEG-3，还没出世就被抛弃了。（MPEG-3 要求传输速率在 20Mbps/sec-40Mbps/sec 间，但这将使画面有轻度扭曲）。除了作为 DVD 的指定标准外，MPEG-2 还可用于为广播，有线电视网，电缆网络以及卫星直播 (DirectBroadcastSatellite)提供广播级的数字视频。

特点：

- 输出码率流速率可适应同步或异步传输，无需固定；
- 适应于逐行或隔行扫描系统；
- 可用于 4:2:0、4:2:2、4:4:4 等亮、色取样；
- 按清晰度可以将图象分为 4 个等级，高级解码器可与低级解码器实现下兼容；
- 按使用的工具和方法不同分为 5 种类型，即允许分层性编码，以保证不同的传输与接收要求。

缺点（视频监控中）：

- 带宽需求比较大：如果要达到比较好的效果，MPEG2 产品需要至少到 6M。而且其带宽只能在 2M、4M、6M、8M 之间切换。
- MPEG2 产品只有 E1 专线的产品，不能是纯 IP 的视讯系统级联。
- MPEG2 产品由于采用的算法比较复杂，决定其编解码过程比较耗时。
- MPEG2 协议当时是为单向广播应用设计的，因此对延时的要求不是很强烈。从目前应用看，会

议经过 MCU 后的延时大概在 1S 左右，因此会议有比较明显的停顿感，如果经过了 MPEG2 和 E1 之间的转换，延时会更高。

相关 RFC（RTP 包）：

a) rfc2038 : RTP Payload Format for MPEG1/MPEG2 Video

<http://tools.ietf.org/rfc/rfc2038.txt>

b) rfc2250 : RTP Payload Format for MPEG1/MPEG2 Video

<http://www.ietf.org/rfc/rfc2250.txt>

1.3 MPEG-4 编码

MPEG-4 标准并非只是 MPEG-2 的替代品，它着眼于不同的应用领域。MPEG-4 的制定初衷主要针对视频会议、可视电话超低比特率压缩（小于 64Kb/s）的需求。在制定过程中，MPEG 组织深深感受到人们对媒体信息，特别是对视频信息的需求由播放型转向基于内容的访问、检索和操作。

MPEG-4 为多媒体数据压缩编码提供了更为广阔的平台，它定义的是一种格式、一种框架，而不是具体算法，它希望建立一种更自由的通信与开发环境。于是 MPEG-4 新的目标就是定义为：支持多种多媒体的应用，特别是多媒体信息基于内容的检索和访问，可根据不同的应用需求，现场配置解码器。编码系统也是开放的，可随时加入新的有效的算法模块。应用范围包括实时视听通信、多媒体通信、远地监测/监视、VOD、家庭购物/娱乐等。

特点：

- 与 MPEG-1 和 MPEG-2 相比，MPEG-4 更适于交互 AV 服务以及远程监控。
- 压缩比高：特别针对低带宽等条件设计算法，因而 MPEG4 的压缩比更高，使低码率的视频传输成为可能
- 图像质量好：MPEG4 的最高图像清晰度为 768*576，远优于 MPEG1 的 352*288，可以达到接近 DVD 的画面效果。这使得它的图像高清晰度非常好。另外，其它的压缩技术由于算法上的局限，在画面中出现快速运动的人或物体和大幅度的场景变化时，图像质量下降。而 MPEG4 采用基于对象的识别编码模式，从而保证良好的清晰度。
- 节省存储空间：同等条件如场景、图像格式和压缩分辨率条件下，经过编码处理的图像文件越小，所占用的存储空间越小。

相关 RFC（RTP 包）：

a) rfc3016 : RTP Payload Format for MPEG-4 Audio/Visual Streams

<http://tools.ietf.org/rfc/rfc3016.txt>

b) rfc3640 : RTP Payload Format for Transport of MPEG-4 Elementary Streams

<http://tools.ietf.org/rfc/rfc3640.txt>

1.4 、Mpeg4 AVC 编码（JVT）

新一代的视频压缩标准。

JVT 是由 ISO/IEC MPEG 和 ITU-T VCEG 成立的联合视频工作组（Joint Video Team），致力于新一代数字视频压缩标准的制定。

JVT 标准在 ISO/IEC 中的正式名称为：MPEG-4 AVC(part10) 标准；在 ITU-T 中的名称：H.264（早期被称

为H.26L)。
其它参考下文的**H.264/AVC**部分。

2、H.23X 系列

(由**ITU/国际电传视讯联盟**/主导, 侧重网络传输, 注意: 只是视频编码)
包括H.261、H.262、H.263、H.263+、H.263++、H.264 (就是MPEG4 AVC-合作的结晶)。

2.1 H.261 编码

H.261又称为P*64, 其中P为64kb/s的取值范围, 是1到30的可变参数, 它最初是针对在ISDN上实现电信会议应用特别是面对面的可视电话和视频会议而设计的。实际的编码算法类似于MPEG算法, 但不能与后者兼容。H.261在实时编码时比MPEG所占用的CPU运算量少得多, 此算法为了优化带宽占用量, 引进了在图像质量与运动幅度之间的平衡折中机制, 也就是说, 剧烈运动的图像比相对静止的图像质量要差。因此这种方法是属于恒定码流可变质量编码而非恒定质量可变码流编码。

优点:

- CPU运算量小: H.261在实时编码时比MPEG所占用的CPU运算量少得多。

缺点:

- 压缩比低
- 缺乏灵活性

相关 RFC(RTP 包):

- a) rfc2032 : RTP Payload Format for H.261 Video Streams
<http://tools.ietf.org/rfc/rfc2032.txt>
- b) rfc4587 : RTP Payload Format for H.261 Video Streams
<http://www.ietf.org/rfc/rfc4587.txt>

2.2 H.262 编码

H.262 是ITU-T的一个数字视频编码标准, 属于视频编解码器。H.262在技术内容上和ISO/IEC的MPEG-2视频标准(正式名称是ISO/IEC 13818-2)一致。

H.262是由ITU-T的VCEG组织和ISO/IEC的MPEG组织联合制定的, 所以制定完成后分别成为了两个组织的标准, 正式名称是"ITU-T建议H.262"和"ISO/IEC 13818-2"。这两个标准在所有的文字叙述上都是相同的(也许除了封面和标价之外)。

PEG2 Video/H.262: **MPEG2(ISO/IEC 13818-2)**, 也被称为**ITU-T H.262**, 是目前消费类电子视频设备中使用最广泛的视频编码标准。MPE2视频用于数字电视广播: 包括陆地, 海底电缆和直接卫星广播。它能在25fps(PAL)或者30fps(NTSC)的固定帧率下达到720x576像素成像。此外, 它也是DVD-V中必需的编解码器。其它内容参考**MPEG2**。

2.3 H.263 编码

H.263标准是甚低码率的图像编码国际标准, 它一方面以H.261为基础, 以混合编码为核心, 其基本原理框图和H.261十分相似, 原始数据和码流组织也相似; 另一方面, H.263也吸收了MPEG等其它一些国际标准中有效、合理的部分, 如: 半像素精度的运动估计、PB帧预测等, 使它性能优于H.261。

H.263使用的位率可小于64Kb/s, 且传输比特率可不固定(变码率)。H.263支持多种分辨率:

SQCIF(128x96)、QCIF、CIF、4CIF、16CIF。

相关 RFC (RTP 包) :

- a) rfc2190 : RTP Payload Format for H.263 Video Streams
<http://tools.ietf.org/rfc/rfc2190.txt>

b) rfc2429 : RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)

<http://tools.ietf.org/rfc/rfc2429.txt>

c) rfc4629 : RTP Payload Format for ITU-T Rec. H.263 Video

<http://tools.ietf.org/rfc/rfc4629.txt>

2.4 H.263+编码

1998年IUT-T推出的H.263+是H.263建议的第2版,它提供了12个新的可协商模式和其他特征,进一步提高了压缩编码性能。如H.263只有5种视频源格式,H.263+允许使用更多的源格式,图像时钟频率也有多种选择,拓宽应用范围;另一重要的改进是可扩展性,它允许多显示率、多速率及多分辨率,增强了视频信息在易误码、易丢包异构网络环境下的传输。另外,H.263+对H.263中的不受限运动矢量模式进行了改进,加上12个新增的可选模式,不仅提高了编码性能,而且增强了应用的灵活性。H.263已经基本上取代了H.261。

2.5 H.263++编码

H263++在H263+基础上增加了3个选项,主要是为了增强码流在恶劣信道上的抗误码性能,同时为了提高增强编码效率。这3个选项为:

选项U——称为增强型参考帧选择,它能够提供更强的编码效率和信道错误再生能力(特别是在包丢失的情形下),需要设计多缓冲区用于存贮多参考帧图像。

选项V——称为数据分片,它能够提供更强的抗误码能力(特别是在传输过程中本地数据被破坏的情况下),通过分离视频码流中DCT的系数头和运动矢量数据,采用可逆编码方式保护运动矢量。

选项W——在H263+的码流中增加补充信息,保证增强型的反向兼容性,附加信息包括:指示采用的定点IDCT、图像信息和信息类型、任意的二进制数据、文本、重复的图像头、交替的场指示、稀疏的参考帧识别。

2.6 H.264 / AVC 编码

H.264集中了以往标准的优点,并吸收了以往标准制定中积累的经验,采用简洁设计,使它比MPEG4更容易推广。H.264创造性地多参考帧、多块类型、整数变换、帧内预测等新的压缩技术,使用了更精细的分像素运动矢量(1/4、1/8)和新一代的环路滤波器,使得压缩性能大大提高,系统更加完善。

H.264 主要有以下几大优点:

- 高效压缩:与H.263+和MPEG4 SP相比,减小50%比特率;
- 延时约束方面有很好的柔性;
- 容错能力;
- 编/解码的复杂性可伸缩性;
- 解码全部细节:没有不匹配;
- 高质量应用;
- 网络友善。

相关 RFC (RTP 包):

a) rfc3984 : RTP Payload Format for H.264 Video

<http://www.ietf.org/rfc/rfc3984.txt>

b) rfc6190 : RTP Payload Format for Scalable Video Coding

<http://www.rfc-editor.org/rfc/rfc6190.txt>

3、其它

3.1 MJPEG 编码

MJPEG全名为"Motion Joint Photographic Experts Group", 是一种视频编码格式, 中文名称翻译为“技术即运动静止图像(或逐帧)压缩技术”。MJPEG广泛应用于非线性编辑领域可精确到帧编辑和多层图像处理, 把运动的视频序列作为连续的静止图像来处理, 这种压缩方式单独完整地压缩每一帧, 在编辑过程中可随机存储每一帧, 可进行精确到帧的编辑, 此外M-JPEG的压缩和解压缩是对称的, 可由相同的硬件和软件实现。但M-JPEG只对帧内的空间冗余进行压缩。不对帧间的时间冗余进行压缩, 故压缩效率不高。采用M-JPEG数字压缩格式, 当压缩比7:1时, 可提供相当于Betacam SP质量图像的节目。

相关 RFC(RTP 包):

- a) Rfc2435 : RTP Payload Format for JPEG-compressed Video
<http://tools.ietf.org/rfc/rfc2435.txt>

3.2 VC-1 编码

VC-1, 全名VC-1视讯编解码器 (Video Codec 1), 是微软所开发的视频编解码系统。VC-1是最后被认可的高清编码格式, 相对于MPEG2, VC-1的压缩比更高, 但相对于H.264而言, 编码解码的计算则要稍小一些, 目前来看, VC-1可能是一个比较好的平衡, 辅以微软的支持, 应该是一只不可忽视的力量。一般来说, VC-1多为“.wmv”后缀, 但这都不是绝对的, 具体的编码格式还是要通过软件来查询。

相关 RFC (RTP 包) :

- a) Rfc4225 : RTP Payload Format for Video Codec 1 (VC-1)
<http://tools.ietf.org/rfc/rfc4225.txt>

3.3 RealVideo 编码

RealVideo格式文件包括后缀名为RA、RM、RAM、RMVB的四种视频格式。Real Video是一种高压缩比的视频格式, 可以使用任何一种常用于多媒体及Web上制作视频的方法来创建RealVideo文件。例如Premiere、VideoShop以及AfterEffects等, 对于文件的播放可用realplayer和暴风影音播放。RealVideo文件是RealNetworks公司开发的一种新型流式视频文件格式, 它包含在RealNetworks公司所制定的音频视频压缩规范RealMedia中, 主要用来在低速率的广域网上实时传输活动视频影像, 可以根据网络数据传输速率的不同而采用不同的压缩比率, 从而实现影像数据的实时传送和实时播放。

3.4 VP8 编码

VP8 是一个开放的影像压缩格式, 最早由 On2 Technologies 开发, 随后由 Google 释出。同时 Google 也释出了 VP8 编码的实现函数库: libvpx, 以 BSD 授权条款的方式发行, 随后也附加了专利使用权。而在经过一些争论之后, 最终 VP8 的授权确认为一个开放源代码授权。

3.5 Theora 编码

Theora 是一个免权利金、开放格式的有损影像压缩技术, 由 Xiph.Org 基金会开发, 该基金会还开发了著名的声音编码技术 Vorbis, 以及多媒体容器档案格式 Ogg。libtheora 则是 Xiph.Org 基金会针对 Theora 格式的实作。Theora 是由 On2 Technologies 公司专属的 VP3 编码器经过开放源代码后衍生而来, 目标是达成比 MPEG-4 Part 2 更好的编码效率。Theora 的命名来自于一个电视节目 Max Headroom 中的角色: Theora Jones。

三、视频编码工具

1、Libav (FFmpeg)

网址: <http://www.libav.org/>

Libav (旧称: FFmpeg) 是一个自由软件, 可以执行音讯和视讯多种格式的录影、转档、串流功能, 包含了 libavcodec — 这是一个用于多个专案中音讯和视讯的解码器函式库, 以及 libavformat —— 一个音讯与视讯格式转换函式库。

1.1 组成元件

ffmpeg: 是一个命令列工具, 用来对视讯档案转换格式, 也支援对电视卡即时编码

ffserver: 是一个 HTTP 多媒体即时广播串流服务器, 支援时光平移

ffplay: 是一个简单的播放器, 基于 SDL 与 FFmpeg 函式库

libavcodec: 包含了全部 FFmpeg 音讯 / 视讯 编解码函式库

libavformat: 包含 demuxers 和 muxer 函式库

libavutil: 包含一些工具函式库

libpostproc: 对于视讯做前处理的函式库

libswscale: 对于影像作缩放的函式库

1.2 主要参数

-i 设定输入档名。

-f 设定输出格式。

-y 若输出档案已存在时则覆盖档案。

-fs 超过指定的档案大小时则结束转换。

-ss 从指定时间开始转换。

-title 设定标题。

-timestamp 设定时间戳。

-vsync 增减 Frame 使影音同步。

影像参数:

-b 设定影像流量, 默认为 200Kbit/秒。(单位请参照下方注意事项)

-r 设定 FrameRate 值, 默认为 25。

-s 设定画面的宽与高。

-aspect 设定画面的比例。

-vn 不处理影像, 于仅针对声音做处理时使用。

-vcodec 设定影像编解码器, 未设定时则使用与输入档案相同之编解码器。

1.3 编码示例

环境:

CentOS 6.2_64 + FFmpeg version 0.6.5

USB 摄像头: 罗技 (Logitech) C110

E-Mail: Mike_Zhang@live.com

mpeg1 视频编码:

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec mpeg1video 1.mpg
```

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec mpeg1video -f rtp rtp://172.16.1.100:7000 > 1.sdp
```

mpeg2 视频编码:

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec mpeg2video 1.mpg
```

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec mpeg2video -f rtp rtp://172.16.1.100:7000 > 1.sdp
```

mpeg4 视频编码:

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec mpeg4 1.mp4
```

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec mpeg4 -f rtp rtp://172.16.1.100:7000 > 1.sdp
```

h261 视频编码:

```
ffmpeg -f video4linux2 -s 176*144 -r 10 -i /dev/video0 -vcodec h261 1.h261
```

h263 视频编码:

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec h263 1.h263
```

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec h263 -f rtp rtp://172.16.1.100:7000 > 1.sdp
```

h264 视频编码:

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec libx264 -subq 5 -g 250 -qmin 10 -qmax 51 -i_qfactor 0.71 1.m4v
```

```
ffmpeg -f video4linux2 -r 30 -i /dev/video0 -vcodec libx264 -subq 5 -g 250 -qmin 10 -qmax 51 -i_qfactor 0.71 -f rtp rtp://172.16.1.100:7000 > 1.sdp
```

2、mencoder

mencoder 是一款命令行方式的视频处理软件，是 Mplayer 自带的编码工具（Mplayer 是 Linux 下的播放器，开源，支持几乎所有视频格式的播放，现在有 windows 和 Mac 版本）。

mencoder 支持几乎所有的格式的视频转换，可以将任意格式转换到任意格式，转换功能可以说是相当强大。目前市面上流行的格式转换器，都是基于 mencoder 开发的 GUI，比如暴风转码，格式工厂等。可以说转换器能办到的，mencoder 都能办到，但 mencoder 能办到的，转换器就不一定能办到了。

3、ffdshow

ffdshow 是一套免费的编解码软件，可支持 H.264、FLV、WMV、MPEG-1、MPEG-2 以及 MPEG-4 等格式视频档。ffdshow 本身并不包含媒体播放器，但各种媒体播放器都可以使用 ffdshow 进行解码。

ffdshow 是一款全能的 DirectShow 解码、编码器，可以解压缩常见的视频格式和几乎所有的音频格式，它还提供了丰富的加工处理选项，可以锐化画面，调节画面的亮度，它还支

持诸多字幕格式, 它能让音频和视频播放更流畅。

4、x264

x264 是一个采用 GPL 授权的视频编码自由软件。x264 的主要功能在于进行 H.264/MPEG-4 AVC 的视频编码, 而不是作为解码器 (decoder) 之用。

自蓝光光碟推出以来, 要制作蓝光光碟产品都需要利用昂贵的软件, 例如索尼的 Blu-code。而 x264 是一款基于 H.264 标准的编码器, 此外它是免费开源的。透过 x264, 蓝光光碟影片就可以存放在不同的媒体中, 例如普通的 DVD 碟片上。只要播放装置集成了适当的解码器, 就可以播放这些蓝光兼容光碟, 而不是只限播放蓝光光碟。

5、Xvid 和 DivX

Xvid (旧称为 XviD) 是一个开放源代码的 MPEG-4 视频编解码器, 它是基于 OpenDivX 而编写的。Xvid 是由一群原 OpenDivX 义务开发者在 OpenDivX 于 2001 年 7 月停止开发后自行开发的。Xvid 支持多种编码模式, 量化 (Quantization) 方式和范围控, 运动侦测 (Motion Search) 和曲线平衡分配 (Curve) 等众多编码技术, 对用户来说功能十分强大。Xvid 的主要竞争对手是 DivX。但 Xvid 是开放源代码的, 而 DivX 则只有免费 (不是自由) 的版本和商用版本。

DivX 是一种将影片的音频由 MP3 来压缩、视频由 MPEG-4 技术来压缩的数字多媒体压缩格式。DivX 由 DivXNetworks 公司发明, DivX 配置 CPU 要求是 300MHz 以上、内存要求是 64M 以上、8M 以上显存的显卡。DivX 视频编码技术是为了打破微软 ASF 的种种协定的束缚, 由 Microsoft mpeg4 v3 修改而来, 使用 MPEG-4 压缩算法。