

基于 FFmpeg + SDL 的视频播放器的制作

(1) 视音频基础知识

雷霄骅

leixiaohua1020@126.com

<http://blog.csdn.net/leixiaohua1020>

目录

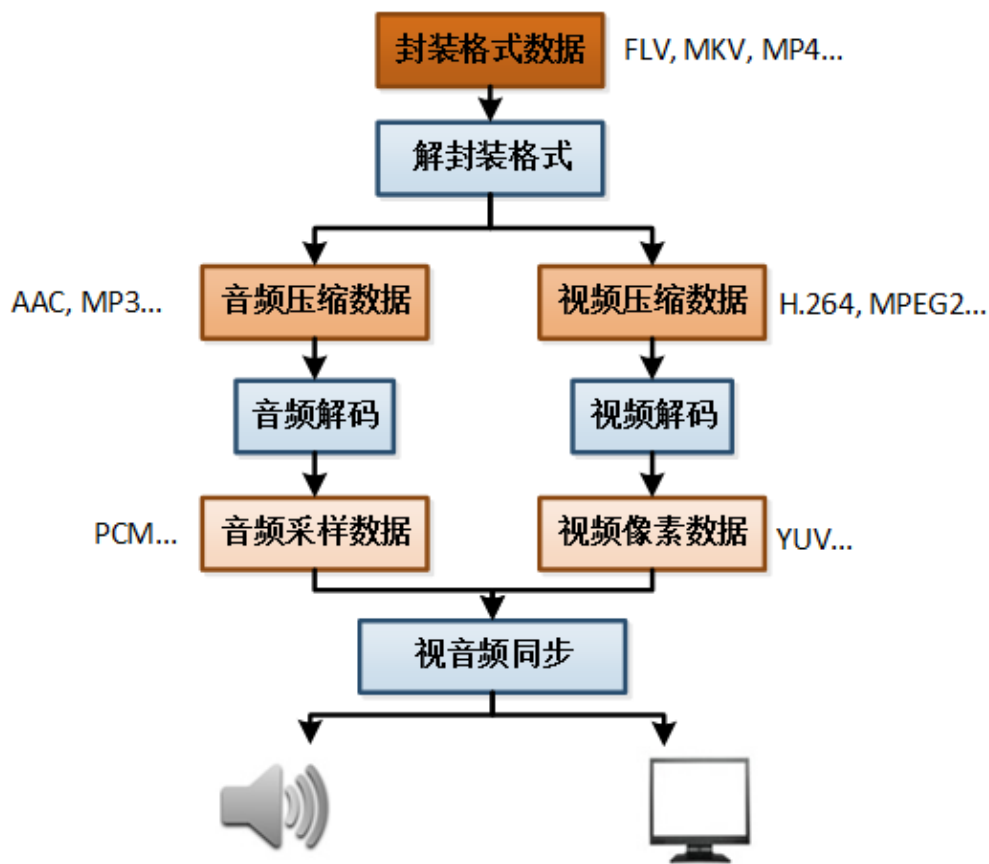
- 前言
- 视频播放器原理
- 封装格式 (MP4 , RMVB , TS , FLV , AVI)
- 视频编码数据 (H.264 , MPEG2 , VC-1)
- 音频编码数据 (AAC , MP3 , AC-3)
- 视频像素数据 (YUV420P , RGB)
- 音频采样数据 (PCM)
- 练习

前言

- 本课程的作用
 - 本课程是视音频技术最基础的入门课程。
 - 本课程的内容是视音频相关行业的必备基础知识。
 - 本课程对未来生活有一定的帮助。
- 学习本课程的前提条件
 - C语言
 - 熟悉Visual C++开发环境

视频播放器原理

播放一个视频文件的流程如下所示。



- 常用播放器

- 跨平台系列（非DirectShow框架）
 - VLC, Mplayer, ffplay……
- Windows系列（ DirectShow框架）
 - 完美解码, 终极解码, 暴风影音……

- 信息查看工具

- 综合信息查看
 - MediaInfo
- 二进制信息查看
 - UltraEdit
- 单项详细分析
 - 封装格式: Elecard Format Analyzer
 - 视频编码数据: Elecard Stream Eye
 - 视频像素数据: YUV Player
 - 音频采样数据: Adobe Audition

- 常用视频播放器



- UltraEdit

The screenshot shows the UltraEdit text editor window titled "[F:\我的教程\小学期\工具\testvideo\中国合伙人.flv] - UltraEdit". The menu bar includes 文件(F), 编辑(E), 搜索(S), 插入(N), 项目(P), 视图(V), 格式(T), 列(L), 宏(M), 脚本(i), 高级(A), 窗口(W), and 帮助(H). The toolbar contains various icons for file operations, editing, and viewing. The main text area displays the hex dump of the file "中国合伙人.flv". The hex dump shows the first 130 hex addresses (00000000h to 00000130h) and their corresponding ASCII values. The first line of the hex dump is highlighted in blue.

```
00000000h: 46 4C 56 01 05 00 00 00 09 00 00 00 00 12 00 01 ; FLV.....
00000010h: 8E 00 00 00 00 00 00 00 02 00 0A 6F 6E 4D 65 74 ; ?.....onMet
00000020h: 61 44 61 74 61 08 00 00 00 0C 00 08 64 75 72 61 ; aData.....dura
00000030h: 74 69 6F 6E 00 40 3E 14 7A E1 47 AE 14 00 05 77 ; tion.@>.z?..w
00000040h: 69 64 74 68 00 40 84 00 00 00 00 00 00 00 06 68 ; idth.@?.....h
00000050h: 65 69 67 68 74 00 40 76 80 00 00 00 00 00 00 0D ; eight.@v?.....
00000060h: 76 69 64 65 6F 64 61 74 61 72 61 74 65 00 40 82 ; videodatarate.@?
00000070h: 4F 80 00 00 00 00 00 09 66 72 61 6D 65 72 61 74 ; O?.....framerat
00000080h: 65 00 40 39 00 00 00 00 00 00 00 0C 76 69 64 65 ; e.@9.....vide
00000090h: 6F 63 6F 64 65 63 69 64 00 40 1C 00 00 00 00 00 ; ocodecid.@.....
000000a0h: 00 00 0D 61 75 64 69 6F 64 61 74 61 72 61 74 65 ; ...audiodatarate
000000b0h: 00 40 57 70 00 00 00 00 00 00 0F 61 75 64 69 6F ; .@Wp.....audio
000000c0h: 73 61 6D 70 6C 65 72 61 74 65 00 40 E5 88 80 00 ; samplerate.@?..
000000d0h: 00 00 00 00 0F 61 75 64 69 6F 73 61 6D 70 6C 65 ; ....audiosample
000000e0h: 73 69 7A 65 00 40 30 00 00 00 00 00 00 00 06 73 ; size.@0.....s
000000f0h: 74 65 72 65 6F 01 01 00 0C 61 75 64 69 6F 63 6F ; tereo....audioco
00000100h: 64 65 63 69 64 00 3F F0 00 00 00 00 00 00 00 0B ; decid.??.....
00000110h: 6D 61 6A 6F 72 5F 62 72 61 6E 64 02 00 04 6D 70 ; major_brand...mp
00000120h: 34 32 00 0D 6D 69 6E 6F 72 5F 76 65 72 73 69 6F ; 42..minor_versio
00000130h: 6E 02 00 01 30 00 11 63 6F 6D 70 61 74 69 62 6C ; n...0..compatibl
```

The status bar at the bottom shows: 按 F1 获取帮助, 位置: 0H, 0, C0, DOS, 936 (ANSI/OEM - 简体中文 GBK), 不高亮, 修改: 20.

- MediaInfo



封装格式

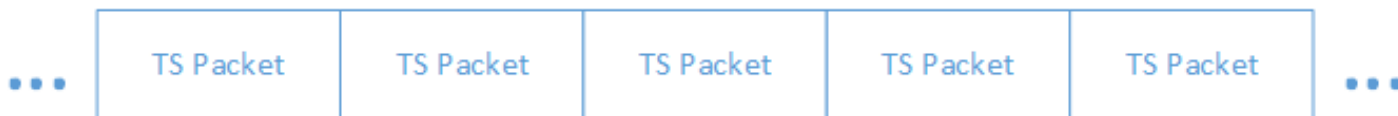
- 封装格式的作用
 - 视频码流和音频码流按照一定的格式存储在一个文件中。
- 封装格式分析工具
 - Elecard Format Analyzer



名称	推出机构	目前使用领域
AVI	Microsoft Inc.	BT下载影视
MP4	MPEG	互联网视频网站
TS	MPEG	IPTV，数字电视
FLV	Adobe Inc.	互联网视频网站
MKV	CoreCodec Inc.	互联网视频网站
RMVB	Real Networks Inc.	BT下载影视

- MPEG2-TS格式简介

- 不包含文件头。数据大小固定（188Byte）的TS Packet构成。



- FLV格式简介

- 包含文件头。数据由大小不固定的Tag构成。



Elecard Video Format Analyzer - D:\cdtworkspace\小学期\工具\testvideo\春晚是什么?.mov

File View Help

Header

Name	Size	Offset
ftyp	20	0x00000000
wide	8	0x00000014
mdat	4630125	0x0000001C
moov	28709	0x0046A689
mvhd	108	0x0046A691
trak	12988	0x0046A6FD
tkhd	92	0x0046A705
edts	36	0x0046A761
elst	28	0x0046A769
mdia	12784	0x0046A785
tapt	68	0x0046D975
trak	15572	0x0046D9B9
udta	33	0x0047168D
5B-□	10578	0x004716AE

Contents

Name	Value
version	0
flags	0x00, 0x00, 0x00
creation_time	0 [1904/1/1 0:00:00]
modification_time	0 [1904/1/1 0:00:00]
timescale	1000
duration	34037 [00:00:34]
rate	1
volume	1
reserved16	0
reserved32	{ 0, 0 }
matrix	{ 65536, 0, 0, 0, 65536, 0, 0, 0, 1073741824 }
pre_defined	{ 0, 0, 0, 0, 0, 0, 0 }
next_track_ID	2

Hex Viewer

Display: hexadecimal Text: ANSI Columns: Auto

0x0046A691 00 00 00 6C 6D 76 68 64 00 00 00 00 00 00 00 00 00 00 00 00 03 E8 00 00 84 F5 ...lmvhd.....è...ö

0x0046A6AD 00 01 00 00 01 00@.....

0x0046A6C9 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 00 40 00 00 00 00 00 00 00 00 00 00 002%

0x0046A6E5 00@.....

0x0046A701 74 72 61 6B 00 00 00 5C 74 6B 68 64 00 00 00 03 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 trak...\tkhd.....

0x0046A71D 00 00 00 00 00 00 84 D0 00@.....

0x0046A739 00@.....

0x0046A755 40 00 00 00 03 50 00 00 01 E0 00 00 00 00 00 24 65 64 74 73 00 00 00 1C 65 6C 73 74 @....P...à....\$edts....elst

0x0046A771 00 00 00 00 00 00 00 01 00 00 84 D0 00 00 04 00 00 01 00 00 00 00 31 F0 6D 64 69 61@.....1&mdia

description FF hex view File Info

Ready

中

视频编码数据

- 视频编码的作用
 - 将视频像素数据（RGB，YUV等）压缩成为视频码流，从而降低视频的数据量。
- 视频编码分析工具
 - Elecard Stream Eye

名称	推出机构	推出时间	目前使用领域
HEVC(H.265)	MPEG/ITU-T	2013	研发中
H.264	MPEG/ITU-T	2003	各个领域
MPEG4	MPEG	2001	不温不火
MPEG2	MPEG	1994	数字电视
VP9	Google	2013	研发中
VP8	Google	2008	不普及
VC-1	Microsoft Inc.	2006	微软平台

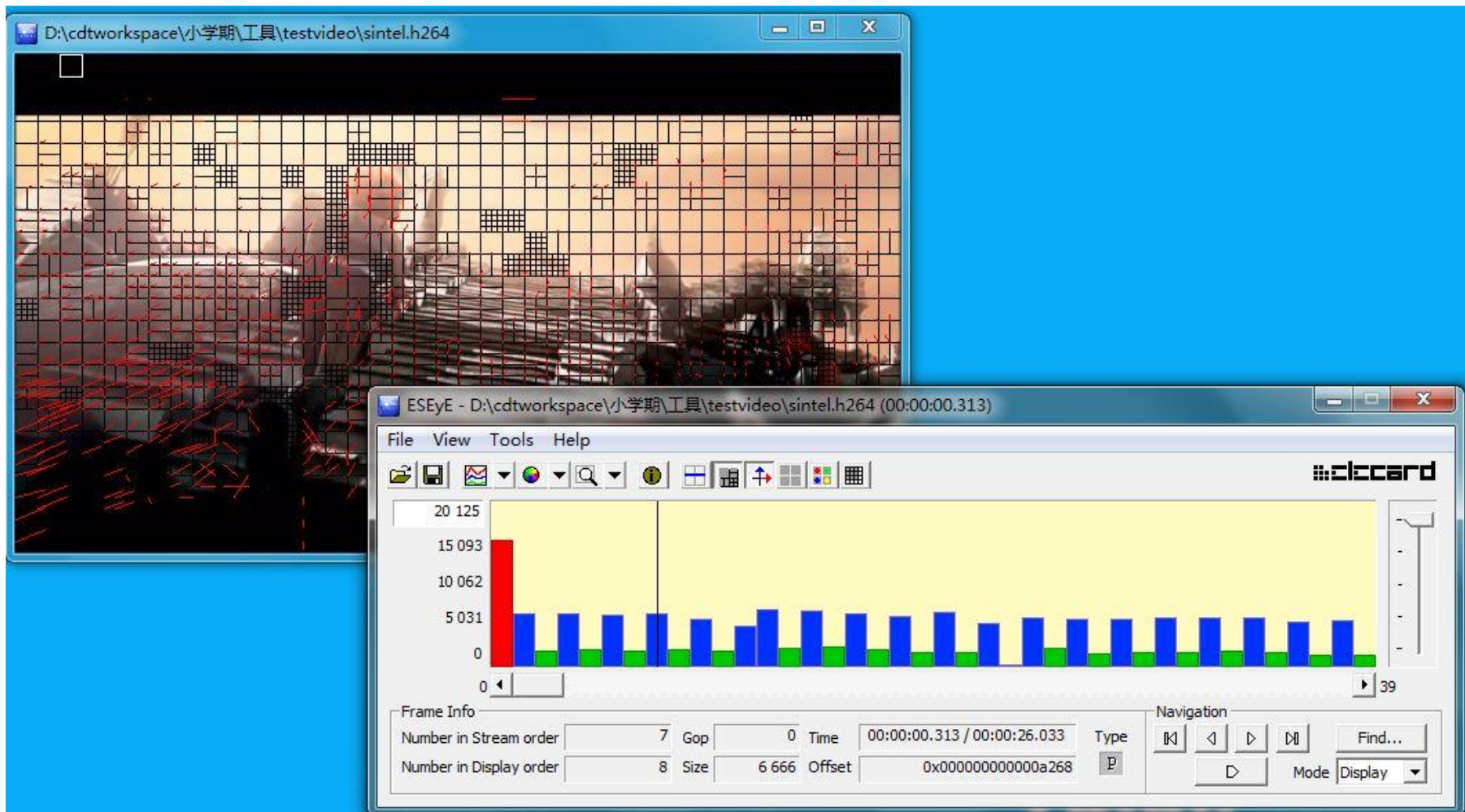
- H. 264格式简介

- 数据由大小不固定的NALU构成
- 最常见的情况下，1个NALU存储了1帧画面的压缩编码后的数据。



- H. 264压缩方法

- 比较复杂。包含了帧内预测、帧间预测、熵编码、环路滤波等环节构成。本课程不对其中的算法做过多介绍。
- 可以将图像数据压缩100倍以上。



红色的帧：I帧；蓝色的帧：P帧；绿色的帧：B帧。

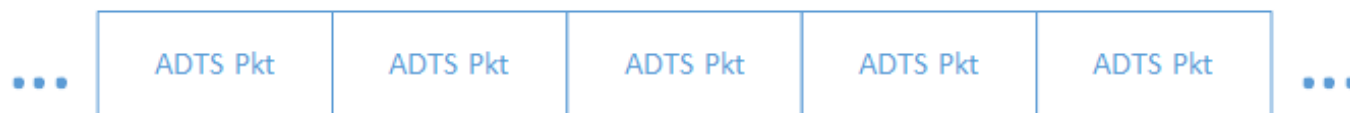
音频编码数据

- 音频编码的作用
 - 将音频采样数据（PCM等）压缩成为音频码流，从而降低音频的数据量。
- 音频编码分析工具
 - 暂不涉及。

名称	推出机构	推出时间	目前使用领域
AAC	MPEG	1997	各个领域（新）
AC-3	Dolby Inc.	1992	电影
MP3	MPEG	1993	各个领域（旧）
WMA	Microsoft Inc.	1999	微软平台

- AAC格式简介

- 数据由大小不固定的ADTS构成



- AAC压缩方法

- 比较复杂。本课程不对其中的算法做过多介绍。
- 可以将音频数据压缩10倍以上。

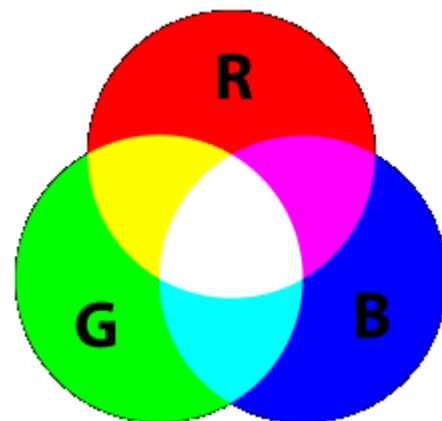
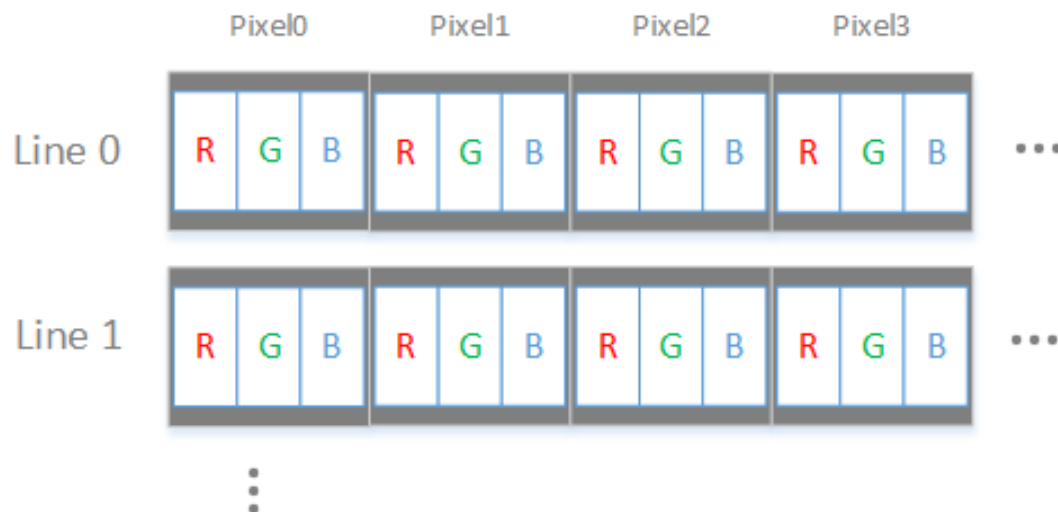
视频像素数据

- 视频像素数据作用
 - 保存了屏幕上每个像素点的像素值。
- 格式
 - 常见的像素数据格式有RGB24, RGB32, YUV420P, YUV422P, YUV444P等。压缩编码中一般使用的是YUV格式的像素数据, 最为常见的格式为YUV420P。
- 特点
 - 视频像素数据体积很大, 一般情况下1小时高清视频的RGB24格式的数据体积为:
$$3600 * 25 * 1920 * 1080 * 3 = 559.9 \text{ GByte}$$

PS: 这里假定帧率为25Hz, 取样精度8bit。
- YUV格式像素数据查看工具
 - YUV Player

• RGB格式简介

- Red、Green、Blue三种颜色，可以混合成世界上所有的颜色。
- 彩色图像中每个点，由R、G、B三个分量组成。
- 以RGB24为例，图像像素数据的存储方式如下：

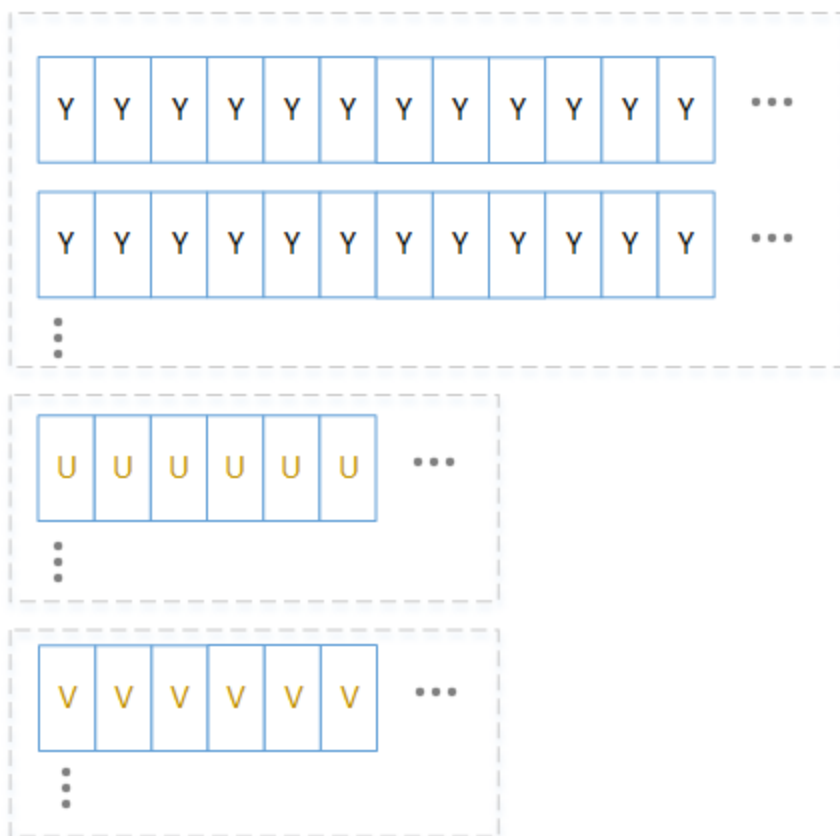


从图中可以看出，RGB24依次存储了每个像素点的R、G、B信息。

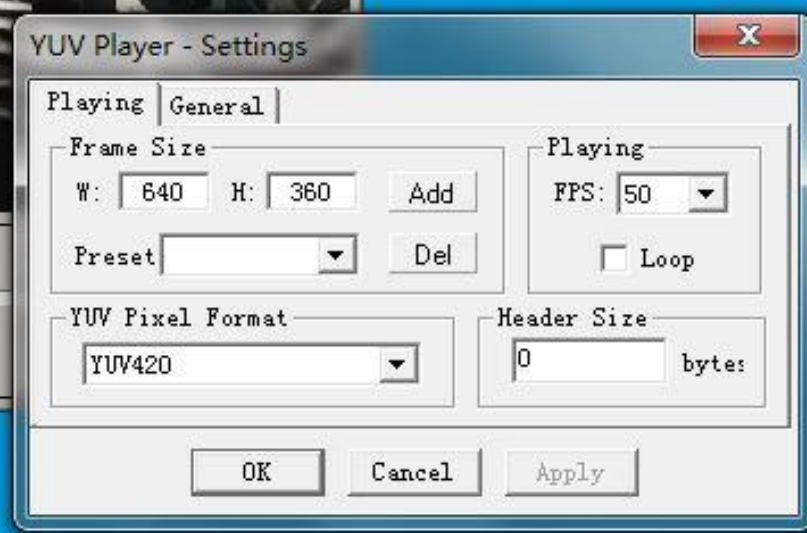
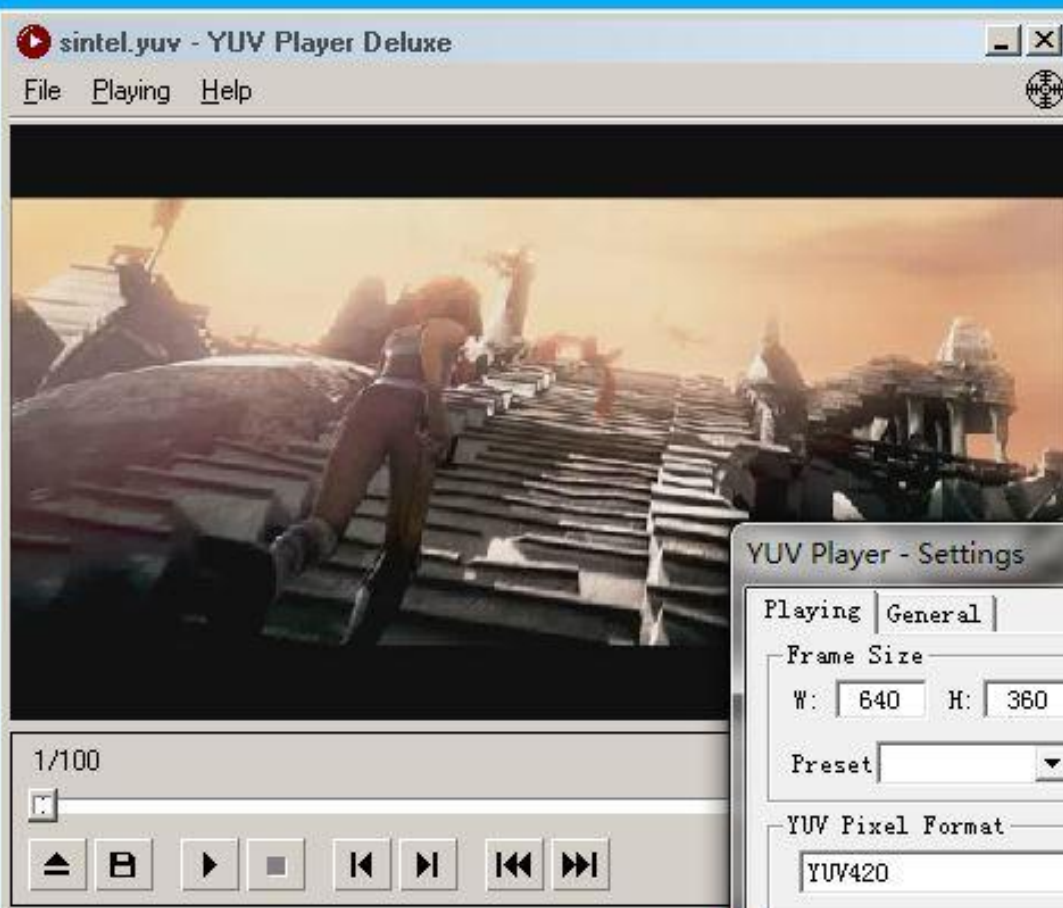
PS：BMP文件中存储的就是RGB格式的像素数据。

• YUV格式简介

- 相关实验表明，人眼对亮度敏感而对色度不敏感。因而可以将亮度信息和色度信息分离，并对色度信息采用更“狠”一点的压缩方案，从而提高压缩效率。



- YUV格式中，Y只包含亮度信息，而UV只包含色度信息。
- 以YUV420P为例，图像像素数据的存储方式如图所示。
- 从图中可以看出，YUV420P首先存储了整张图像的Y信息，然后存储整张图像的U信息，最后存储了整张图像的V信息。



音频采样数据

- 音频采样数据作用
 - 保存了音频中每个采样点的值。
- 特点
 - 音频采样数据体积很大，一般情况下一首4分钟的PCM格式的歌曲体积为：

$$4*60*44100*2*2=42.3\text{MByte}$$

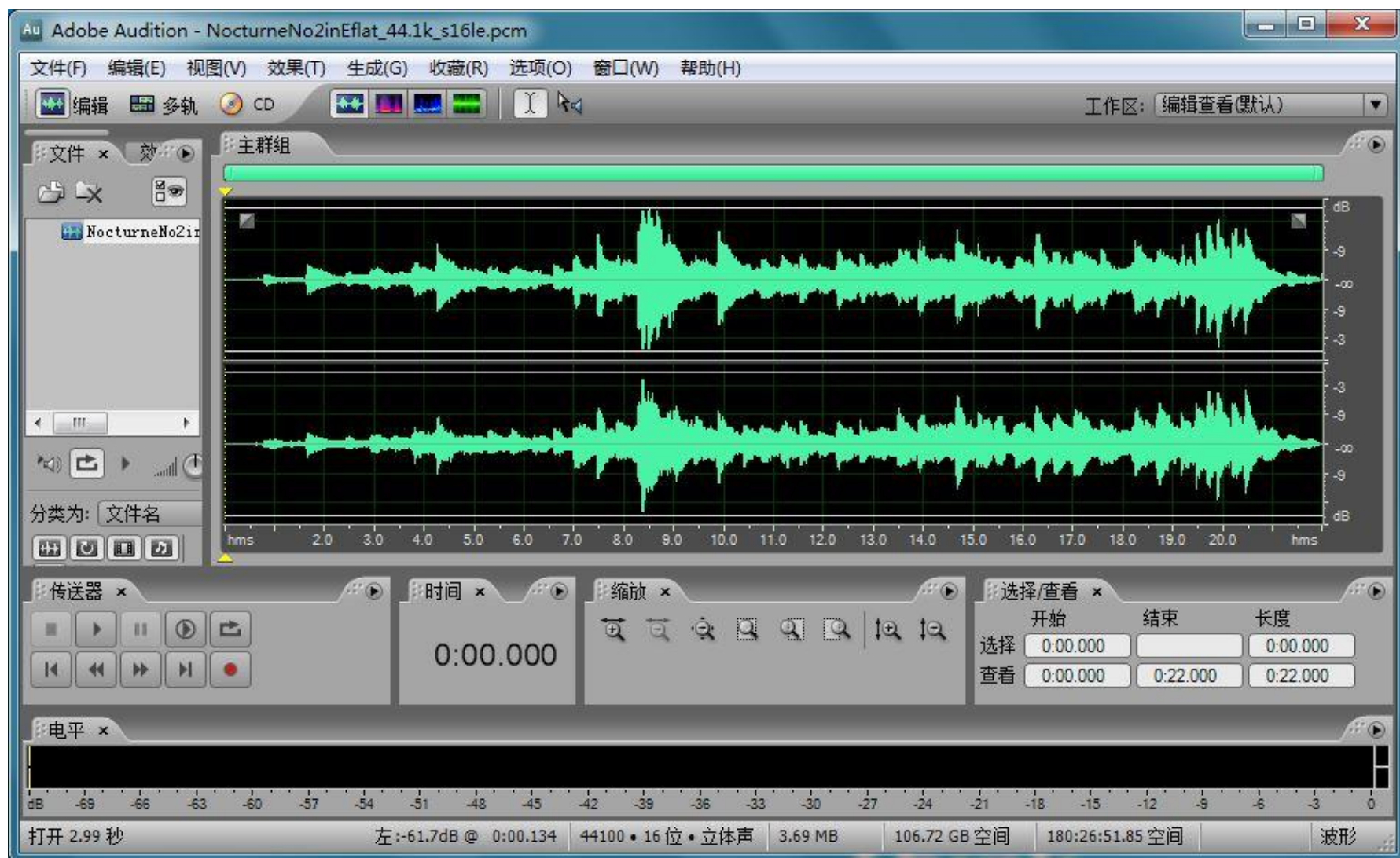
PS: 这里假定采样率为44100Hz，采样精度为16bit。

- 音频采样数据查看工具：Adobe Audition

- PCM格式简介

- 单声道的情况下按照顺序存储每个采样点的数据。
- 双声道的情况下按照“左右、左右”的顺序存储每个采样点两个声道的数据。





实验

- 对视频文件进行分析，对下列信息进行截图：
 - 任意4个文件的总体信息（MediaInfo）
 - MKV、MP4、AVI封装格式信息（Elecard Format Analyzer）
 - H.264码流分析信息（Elecard Stream Eye）
 - YUV视频像素数据（YUV Player）
 - PCM音频采样数据（Audition）