

第 3 章



音视频开发常用工具

工欲善其事，必先利其器。音视频知识比较复杂，掌握几款常用的音视频工具，对于初学者非常有帮助，可以做到事半功倍。VLC 是一款功能很强大的开源播放器，支持多种常见音视频格式，支持多种流媒体传输协议，也可当作本地流媒体服务器使用。MediaInfo 用来分析视频和音频文件的编码和内容信息。Elecard Stream Analyzer 是一款简单小巧的码流分析工具，通过该软件，用户可以快速地分析视频序列码流。FFmpeg 是一个跨平台的音视频处理库，是一套可以用来记录、转换数字音视频，并能将其转化为流的开源计算机程序。

3.1 VLC 播放器简介

VLC 是一款功能很强大的开源播放器，VLC 的全名为 Video Lan Client，是一个开源的、跨平台的视频播放器。VLC 支持多种常见音视频格式，支持多种流媒体传输协议，也可当作本地流媒体服务器使用。官网下载网址为 <https://www.videolan.org/>。

3.1.1 VLC 播放器

VLC 多媒体播放器是 VideoLAN 计划的多媒体播放器。它支持众多音频与视频解码器及文件格式，并支持 DVD 影音光盘、VCD 影音光盘及各类流式协议。它也能作为 unicast 或 multicast 的流式服务器在 IPv4 或 IPv6 的高速网络连接下使用。它融合了 FFmpeg 的解码器与 libdvdcss 程序库，使其有播放多媒体文件及加密 DVD 影碟的功能。作为音视频的初学者，很有必要熟练掌握 VLC 这个工具。

3.1.2 VLC 的功能列表

VLC 是一款自由、开源的跨平台多媒体播放器及框架，可播放大多数多媒体文件、DVD、CD、VCD 及各类流媒体协议文件。VLC 支持大量的音视频传输、封装和编码格式，下面列出简要的功能列表。

- (1) 操作系统包括 Windows、Windows CE、Linux、Mac OS X、BEOS、BSD 等。
- (2) 访问形式包括文件、DVD/VCD/CD、HTTP、FTP、TCP、UDP、HLS、RTSP 等。
- (3) 编码格式包括 MPEG、DIVX、WMV、MOV、3GP、FLV、H. 264、FLAC 等。

- (4) 视频字幕包括 DVD、DVB、Text、Vobsub 等。
- (5) 视频输出包括 DirectX、X11、XVideo、SDL、FrameBuffer、ASCII 等。
- (6) 控制界面包括 WxWidgets、QT、Web、Telnet、Command line 等。
- (7) 浏览器插件包括 ActiveX、Mozilla 等。

3.1.3 VLC 播放网络串流

VLC 播放一个视频大致分为 4 个步骤：第一步，access，即从不同的源获取流；第二步，demux，即把通常合在一起的音频和视频分离（有的视频也包含字幕）；第三步，decode，即解码，包括音频和视频的解码；第四步，output，即输出，也分为音频和视频的输出（aout 和 vout）。

使用 VLC 可以很方便地打开网络串流。首先单击主菜单的“媒体”，选择“打开网络串流”，如图 3-1 所示，然后在弹出的对话框界面中输入网络 URL，如图 3-2 所示，单击“播放”按钮，即可看到播放的网络流效果，如图 3-3 所示。测试地址为 CCTV-1 高清频道 <http://ivi.bupt.edu.cn/hls/cctv1hd.m3u8>。

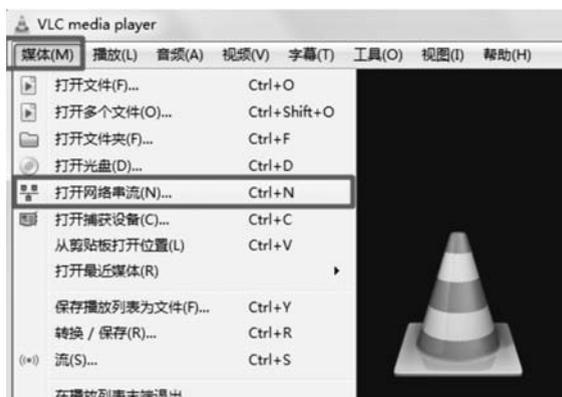


图 3-1 VLC 打开网络串流



图 3-2 VLC 输入网络串流地址



图 3-3 VLC 播放 CCTV-1 高清频道

3.1.4 VLC 作为流媒体服务器

VLC 的功能很强大,它不仅是一个视频播放器,也可作为小型的视频服务器,一边播放一边转码,把视频流发送到网络上。VLC 作为视频服务器的具体步骤如下。

- (1) 单击主菜单“媒体”中的“流”选项。
- (2) 在弹出的对话框中单击“添加”按钮,选择一个本地视频文件,如图 3-4 所示。



图 3-4 VLC 流媒体服务器之打开本地文件

- (3) 单击页面下方的“串流”下拉列表框中的“串流”选项,添加串流协议,如图 3-5 所示。
- (4) 该页面会显示刚才选择的本地视频文件,然后单击“下一个”按钮,如图 3-6 所示。



图 3-5 VLC 流媒体服务器之添加串流协议



图 3-6 VLC 流媒体服务器之文件来源

(5) 在该页面单击“添加”按钮，选择具体的流协议，例如 RTSP，然后单击“下一个”按钮，如图 3-7 所示。

(6) 在该页面的下拉列表中选择“Video-H. 264+MP3(TS)”，然后单击“下一个”按钮，如图 3-8 所示。

注意：一定要选中左上方的“激活转码”，并且需要是 TS 流格式。

(7) 在该页面可以看到 VLC 生成的所有串流输出参数，然后单击“流”按钮即可，如图 3-9 所示。



图 3-7 VLC 流媒体服务器之选择 RTSP 协议

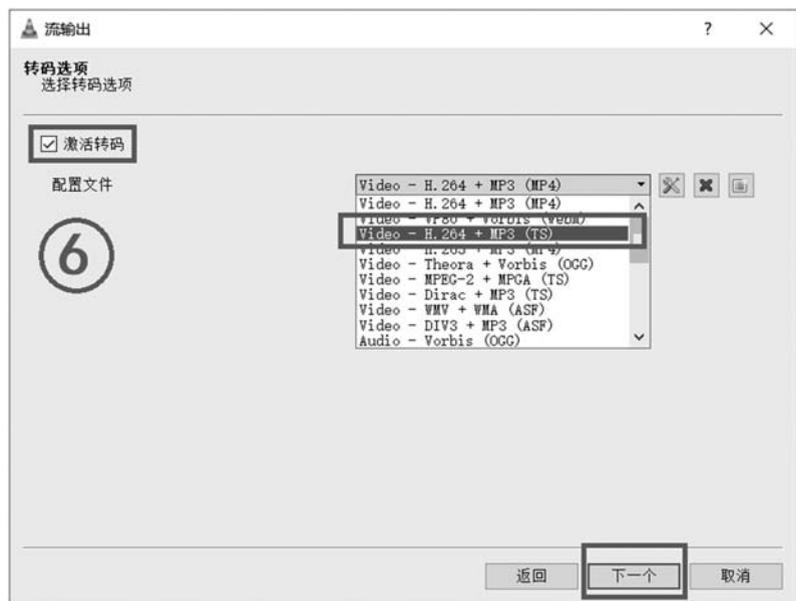


图 3-8 VLC 流媒体服务器之 Video-H.264+MP3(TS)



图 3-9 VLC 流媒体服务器之串流输出参数字符串

3.2 MediaInfo 简介

3.2.1 MediaInfo

MediaInfo 用来分析视频和音频文件的编码和内容信息,是一款自由软件,可以免费使用、免费获得源代码,许可协议是 GNU GPL/LGPL。

1. 获取多媒体文件信息

MediaInfo 可以获取的多媒体文件的基本信息,具体包括以下几方面。

- (1) 内容信息: 标题、作者、专辑名、音轨号、日期、总时间等。
- (2) 视频: 编码器、宽高比、帧频率、码率、比特率等。
- (3) 音频: 编码器、采样率、声道数、语言、比特率等。
- (4) 文本: 语言、字幕等。
- (5) 段落: 段落数、列表等。

2. 支持的格式

MediaInfo 支持的格式很多,具体包括以下几种格式。

- (1) 视频: MKV、OGM、AVI、DivX、WMV、QuickTime、MPEG、DVD 等。
- (2) 编码器: DivX、XviD、MSMPEG-4、ASP、H. 264、AVC 等。
- (3) 音频: OGG、MP3、WAV、RA、AC-3、DTS、AAC、M4A、AU、AIFF 等。
- (4) 字幕: SRT、SSA、ASS、SAMI 等。

3. 查看方式

MediaInfo 支持众多查看方式,包括文本、表格、树形图、网页、HTML、XML 等,如图 3-10 所示。它支持 3 种发布版本,包括图形界面、命令行、DLL(动态链接库),还可以与 Windows 资源管理器整合,包括拖放、右击菜单等。

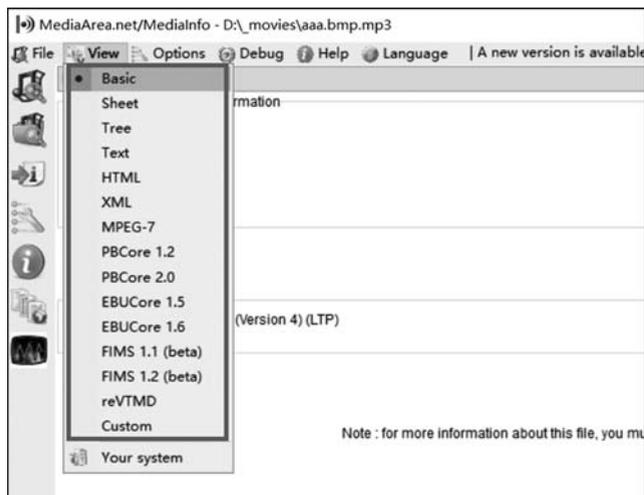


图 3-10 MediaInfo 查看方式

4. 国际化

MediaInfo 的软件界面可以轻松实现本地化,但需要当地的志愿者翻译语言文件。

3.2.2 MediaInfo 使用方法

如果需要查看少数媒体文件的信息,则可以直接把文件拖入 MediaInfo 应用界面,然后就会直接显示出文件的相关信息,也可以直接查看整个文件夹下的媒体文件信息,可以单击界面左侧中间的图标。无论采用哪种方式,都可以查看视频、音频、图片的格式信息,如图 3-11 所示。

General	
MPEG-4 (Base Media): 1.86 GiB, 2 h 16 min	Overall bit rate mode: Variable
1 Video stream: AVC	Overall bit rate: 1 963 kb/s
1 Audio stream: AAC	Encoded date: UTC 2020-01-20 13:23:01
	Tagged date: UTC 2020-01-20 13:23:01
Go to the website of a player for this file	
Video	
1 863 kb/s, 1920*1080 (16:9), at 24.000 FPS, AVC (High@L3.1) (CABAC / 4 Ref Frames)	
Go to the website of this codec	
Audio	
96.0 kb/s, 48.0 kHz, 2 channels, AAC (LC)	

图 3-11 MediaInfo 显示媒体信息

3.2.3 MediaInfo 参数说明

媒体文件信息涉及的内容比较复杂,这里简单介绍一下。此处先把视图切换为 HTML (View 菜单 → HTML)。可以看出,媒体的详细信息主要包括 3 部分的参数,分别是 General、Video、Audio,分别如图 3-12 和图 3-13 所示。

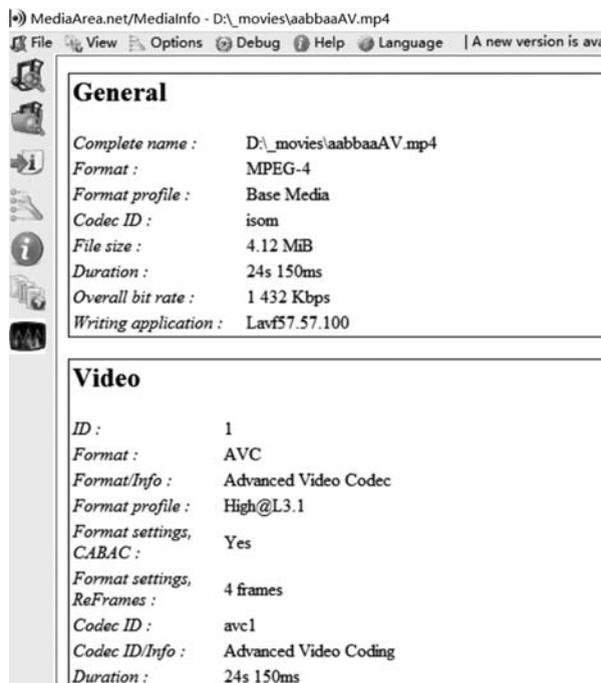


图 3-12 MediaInfo 以 HTML 格式显示媒体详细信息之一

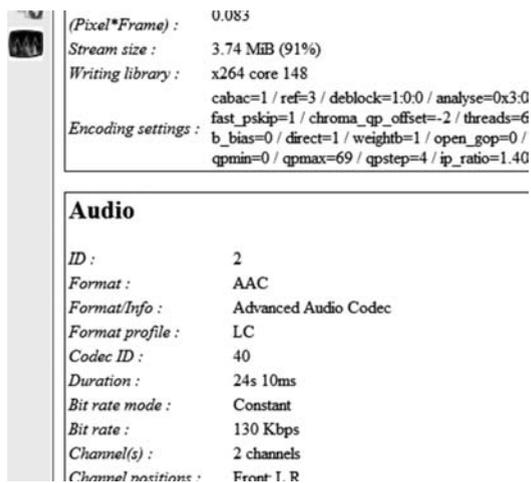


图 3-13 MediaInfo 以 HTML 格式显示媒体详细信息之二

(1) General 主要包括视频封装格式的信息,包括文件大小、文件时长、比特率、编码时间等。

(2) Video 主要包括视频编码的相关信息,包括编码器、Profile&Level、是否使用算术熵编码、比特率、视频文件大小、视频尺寸、帧率模式、帧率、色彩空间、扫描类型(逐行/隔行)、编码设置等。

(3) Audio 主要包括音频编码的相关信息,包括格式、声道数、编码格式、Profile、时长、比特率、是否有损压缩、音频的帧率等信息。

3.3 FlvAnalyser 简介

3.3.1 FLV 简介

直播推流的时候需要用到 RTMP 的视频数据格式。RTMP 的视频格式和 FLV 相似,通过查看 FLV 的格式文档,可以通过分析 FLV 格式来解析 RTMP 格式。RTMP 中的数据就是由 FLV 的 TAG 中的数据区构成的。FLV 是流媒体封装格式,可以将其数据看为二进制字节流。

总体上看,FLV 包括文件头(File Header)和文件体(File Body)两部分,其中文件体由一系列的 Tag 及 Tag Size 对组成,如图 3-14 所示。

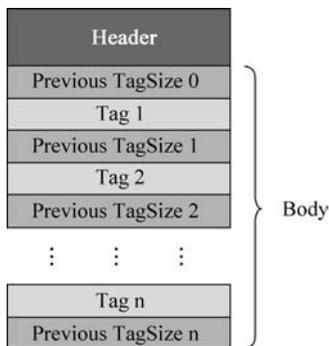


图 3-14 FLV 格式结构图

3.3.2 FlvAnalyser

FlvAnalyser 非常简单方便,功能也很强大,包括文件格式分析、数据分析、十六进制分析、时间戳分析、码率分析、音视频同步分析、日志记录、语法指南(FLV 基本语法)、视频或音频 ES 提取文件、时间信息提取等。

查看文件信息,单击 AV 工具栏按钮,选择一个本地 FLV 文件,即可看到分析的效果,如图 3-15 所示。

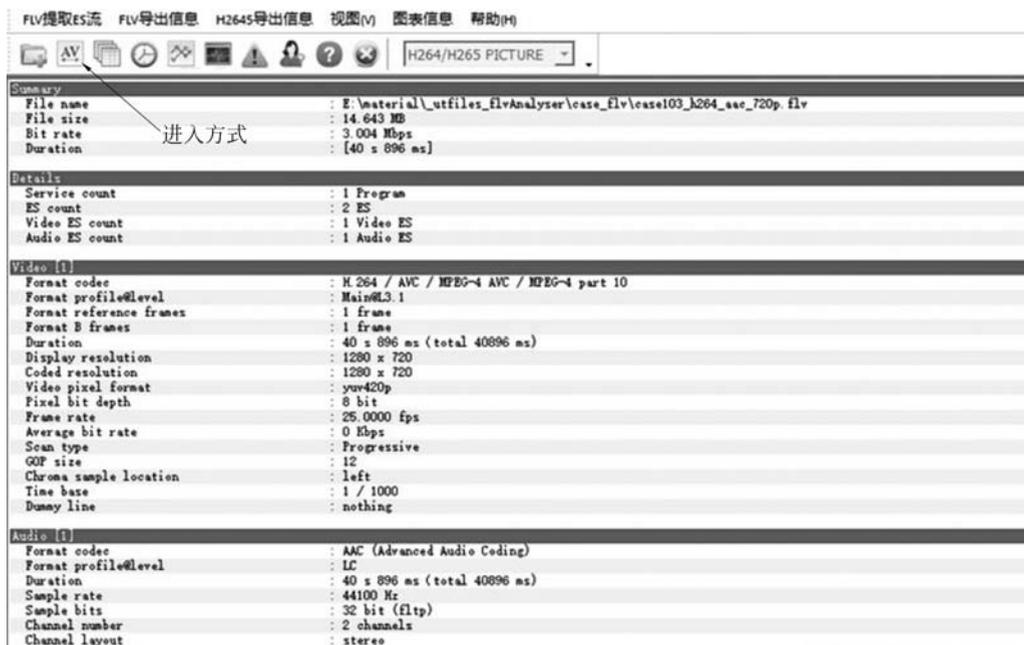


图 3-15 FlvAnalyser 分析文件信息

查看 Tag 列表信息,单击“FLV 提取 ES 流”工具栏按钮,可以查看语法详情、NALU 标注、二进制与十六进制切换等,显示效果如图 3-16 所示。

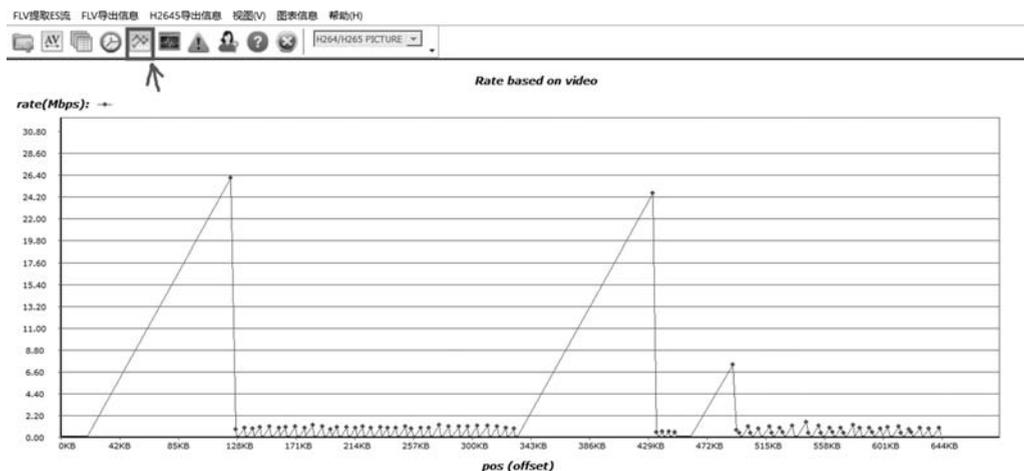


图 3-16 FlvAnalyser 分析 Tag 列表信息

查看时间信息,单击“小时钟”工具栏按钮,可以查看时间信息,显示效果如图 3-17 所示。

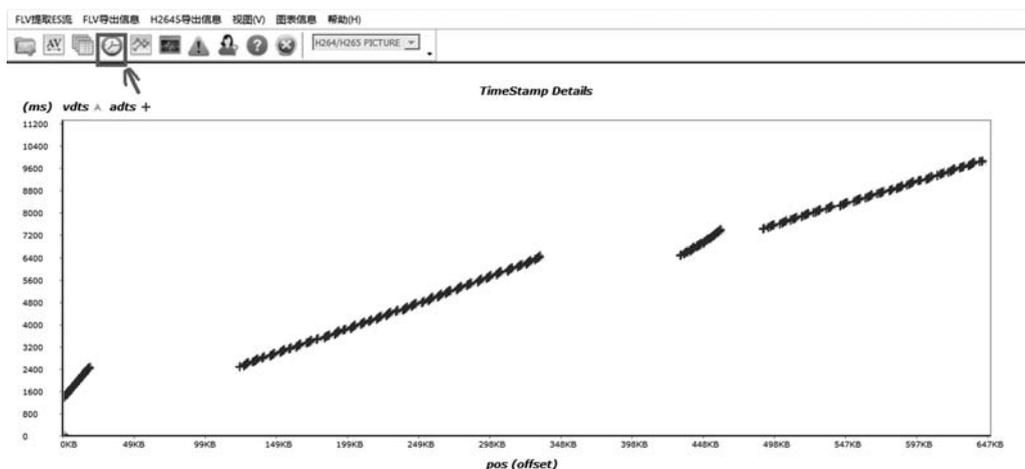


图 3-17 FlvAnalyser 分析时间信息

3.4 Elecard Stream Analyzer 码流分析工具

Elecard Stream Analyzer 是一款简单小巧的码流分析工具,通过该软件,用户可以快速地分析视频序列码流。软件操作简单,使用方便,用户只需将视频文件导入软件内,系统就会自动分析文件,分析后就会显示视频码的文件大小、码流类型、数据包数等内容,方便用户对视频的质量进行初步评估,可有效地改善视频的拍摄质量及制定相应的修改方案。

3.4.1 简介

码流是指视频文件在单位时间内使用的数据流量,是视频编码中画面质量控制中最重要的部分。在同样的分辨率下,视频文件的码流越大,压缩比就越小,画面质量就越好。

多码流技术是通过在编码过程中同时产生多种不同码流及分辨率的流媒体数据,根据用户实际网络带宽条件为之自动分配相对最佳解码画质的解决方案。在实际网络直播应用中,由于位于不同网络位置的访问者所在网络环境存在差异,而仅以某种固定码流分辨率进行网络直播流媒体传送往往会导致网速较高的用户看到的画质仍不够清晰,网速较低的用户由于解码时间过长而使画面不够流畅,为解决二者的矛盾使访问者浏览到尽可能兼顾清晰和流畅的直播内容,采用多码流技术成为一种最简单、最有效的办法。

Elecard StreamEye Tools 是一款分析音视频的好工具,包括 Elecard Stream Analyzer、Elecard StreamEye、Elecard YUV Viewer。

(1) Elecard StreamEye 用于编码视频的可视化表现,以及流结构分析,这些流是 MPEG-1/2/4 或 AVC/H.264 视频基本流 (Video Element Stream, VES)、MPEG-1 的系统流 (System Stream, SS)、MPEG-2 的程序流 (Program Stream, PS)、MPEG-2 的传输流 (Transport Stream, TS)。

(2) Elecard YUV Viewer 是用来查看 YUV 视频文件数据序列的,和其他的文件相比较,找到二进制的图像是否是匹配的,并且可观看比较结果。应用程序允许用户计算度量的质量,例如 PSNR、NQM 和 VQM。

(3) Elecard Stream Analyzer 用于编码媒体流的语法分析,以及人类可读形式的展示。可以操作 MPEG-1 Video/Audio、MPEG-2 Video/Audio、AAC、AC-3、AVC/H. 264 等。Elecard Stream Analyzer 的主界面如图 3-18 所示。

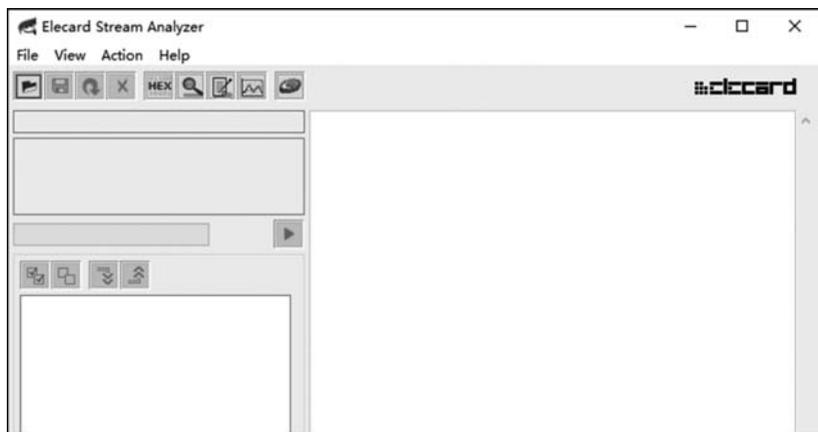


图 3-18 Elecard Stream Analyzer 主界面

3.4.2 功能列表

Elecard Stream Analyzer 是一个专业的视频码流分析工具,功能非常强大,拥有码流错误报告、TS 错误侦测和基于 ETSI TR101-290 的分析功能。能够对编码器媒体流进行深入句法分析;支持 H. 264/AVC、MPEG-2 TS/ PS 等多种常用的视频格式;可以将 ES 码流导出到文件;支持 HEX 浏览、导航和搜索;支持音视频交织分析;支持自动分析的命令行模式等。

Elecard Stream Analyzer 所支持的视频格式主要包括 MPEG-2 TS、MPEG-2 PS、MPEG-1/2 Video Stream、AVC/H. 264 Video Stream、HEVC/H. 265 Video Stream、VP9 Video Stream、MP4 File Container、MKV File Container、AVI File Container 等。

Elecard Stream Analyzer 支持的音频格式主要包括 MPEG-1/2 Audio Layer 1/2/3、Dolby Digital Audio、高级音频编码(Advanced Audio Coding, AAC)等。

安装好软件之后,打开一个本地视频文件,主界面如图 3-19 所示,会显示 SEI、序列参数集(Sequence Parameter Set, SPS)、图像参数集(Picture Parameter Set, PPS)、I/P/B Slice 等信息。

单击某帧可以查看详细信息,如图 3-20 所示,包括 frame_num、slice_type 等。

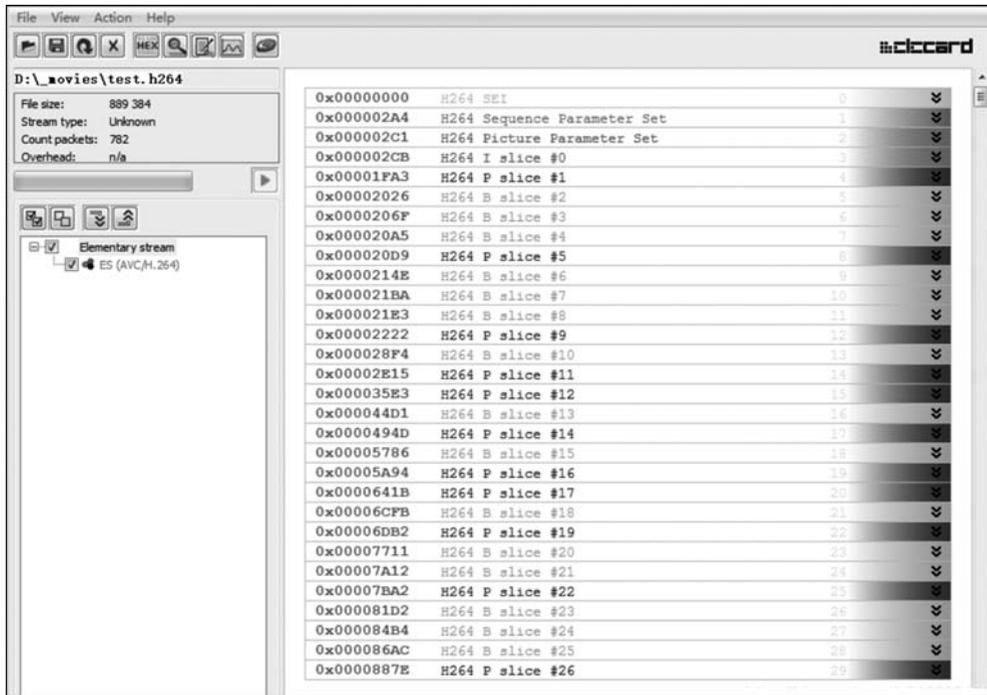


图 3-19 Elecard Stream Analyzer 打开本地 H.264 文件

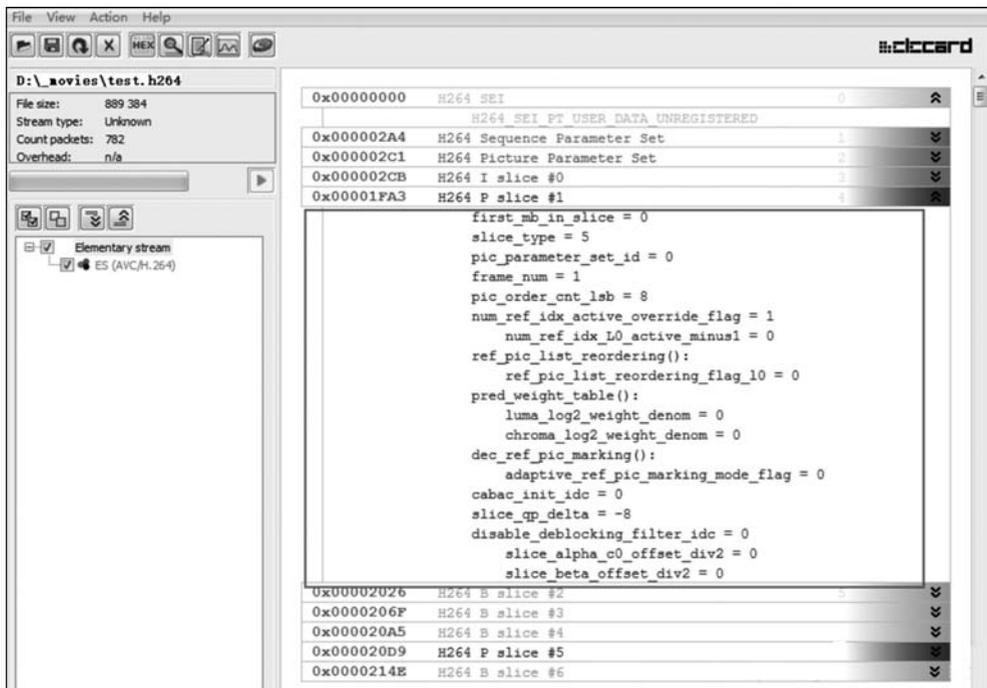


图 3-20 Elecard Stream Analyzer 查看帧详细信息

3.5 FFMpeg 简介

3.5.1 FFMpeg

FFmpeg 是一个跨平台的音视频处理库,是一套可以用来记录、转换数字音视频,并能将其转化为流的开源计算机程序,支持 Windows、Linux、Mac 等。FFmpeg 采用 LGPL 或 GPL 许可证,提供了录制、转换及流化音视频的完整解决方案,包含非常先进的音视频编解码库 libavcodec。为了保证高可移植性和编解码质量,libavcodec 中很多编解码算法都是全新开发的。

FFmpeg 包括几个常用的模块库,如下所述。

(1) libavformat: 用于各种音视频封装格式的生成和解析,包括获取解码所需信息以生成解码上下文结构。

(2) libavcodec: 用于各种类型音视频编解码。

(3) libavutil: 包含一些公共的工具函数。

(4) libswscale: 用于视频场景比例缩放、色彩映射转换等。

(5) libpostproc: 用于后期效果处理等。

(6) libswresample: 提供音频重采样功能,包括采样频率、声道格式等。

(7) libavfilter: 用于滤波器处理,如音视频倍速、水平翻转、叠加文字等功能。

(8) libavdevice: 包含输入/输出设备的库,实现音视频数据的抓取或渲染。

FFmpeg 提供了 3 个命令行工具,如下所述。

(1) ffmpeg: 编解码小工具,可用于格式转换、解码或电视卡即时编码等。

(2) ffmpegserver: 一个 HTTP 多媒体即时广播串流服务器。

(3) ffmpegplay: 一个简单的播放器,使用 ffmpeg 库解析和解码,通过 SDL 显示。

FFmpeg 的开发分为两种,一种方式是直接使用所提供的这 3 个命令行工具进行多媒体处理,另一种是使用封装的上述 8 个模块库进行二次开发。

3.5.2 FFMpeg 命令行

FFmpeg 提供的命令行功能非常强大,包括但不限于以下功能。

(1) 列出支持的格式。

(2) 剪切一段媒体文件。

(3) 提取一个视频文件中的音频文件。

(4) 从 MP4 文件中抽取视频流并可导出裸的 H.264 数据。

(5) 视频静音,即只保留视频,使用命令参数-an。

- (6) 使用 AAC 音频数据和 H.264 视频生成 MP4 文件。
- (7) 音频格式转换。
- (8) 从 WAV 音频文件中导出 PCM 裸数据。
- (9) 将一个 MP4 文件转换为一个 GIF 动图。
- (10) 使用一组图片生成 GIF 动图。
- (11) 淡入效果器使用。
- (12) 淡出效果器使用。
- (13) 将两路声音合并,例如加背景音乐。
- (14) 为视频添加水印效果。
- (15) 视频提亮效果器。
- (16) 视频旋转效果器的使用。
- (17) 视频裁剪效果器的使用。
- (18) 将一段视频推送到流媒体服务器上。
- (19) 将流媒体服务器上的流下载并保存到本地计算机。
- (20) 将两个音频文件以两路流的形式封装到一个文件中。

注意: 本书仅提供命令行功能列表,更详细的参数使用可关注后续的图书。

3.5.3 FFmpeg 开发包

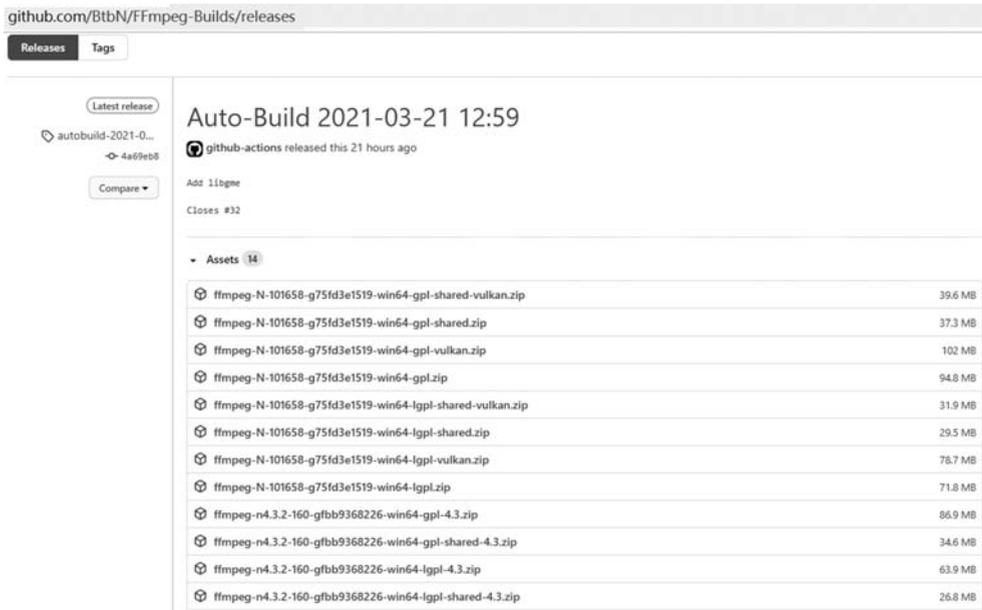
当前 FFmpeg 最新的版本为 4.3.1,包括静态库、动态库、开发者 3 种版本。

(1) Static(静态库版本): 其中只有 3 个应用程序,包括 ffmpeg.exe、ffplay.exe、ffprobe.exe,每个 exe 应用程序的体积都很大,相关的 dll 文件已经被编译到 exe 程序中了。作为工具而言此版本最合适,不依赖动态库,为单个可运行程序。

(2) Shared(动态库版本): 其中除了 3 个应用程序 ffmpeg.exe、ffplay.exe、ffprobe.exe 之外,还有一些动态 dll 文件,例如 avcodec-54.dll 之类的文件。Shared 中的 exe 程序体积很小,它们在运行的时候,到相应的 dll 文件中调用功能。程序运行过程必须依赖于提供的 dll 文件,开发程序时必须下载该版本,因为只有该版本中有 dll 动态库,需要注意 Dev(开发者版本)中不包含这些 dll 动态库。

(3) Dev(开发者版本): 此版本用于开发,其中包含了库文件 xxx.lib 及头文件 xxx.h,这个版本不包含 exe 文件和 dll 文件。Dev 版本中 include 文件夹内包含了所有头文件,lib 文件夹中包含了所有编译开发所需要的库,但没有运行库,所以需要从 Shared 版本中获取。

最新的下载网址为 <https://github.com/BtbN/FFmpeg-Builds/releases>,如图 3-21 所示。也可扫描本书前言处“本书源代码下载”二维码来下载开发包和源代码。下载之后,解压出来即可,开发环境可以选择 Windows 10 64bit + QT5.9.8/VIS2019。



github.com/BtbN/FFmpeg-Builds/releases

Releases Tags

Latest release

autobuild-2021-0...
4a697eb6

Compare

Auto-Build 2021-03-21 12:59

github-actions released this 21 hours ago

Add 11bgee

Closes #32

Assets 14

ffmpeg-N-101658-g75fd3e1519-win64-gpl-shared-vulkan.zip	39.6 MB
ffmpeg-N-101658-g75fd3e1519-win64-gpl-shared.zip	37.3 MB
ffmpeg-N-101658-g75fd3e1519-win64-gpl-vulkan.zip	102 MB
ffmpeg-N-101658-g75fd3e1519-win64-gpl.zip	94.8 MB
ffmpeg-N-101658-g75fd3e1519-win64-igpl-shared-vulkan.zip	31.9 MB
ffmpeg-N-101658-g75fd3e1519-win64-igpl-shared.zip	29.5 MB
ffmpeg-N-101658-g75fd3e1519-win64-igpl-vulkan.zip	78.7 MB
ffmpeg-N-101658-g75fd3e1519-win64-igpl.zip	71.8 MB
ffmpeg-n4.3.2-160-gfbb9368226-win64-gpl-4.3.zip	86.9 MB
ffmpeg-n4.3.2-160-gfbb9368226-win64-gpl-shared-4.3.zip	34.6 MB
ffmpeg-n4.3.2-160-gfbb9368226-win64-igpl-4.3.zip	63.9 MB
ffmpeg-n4.3.2-160-gfbb9368226-win64-igpl-shared-4.3.zip	26.8 MB

图 3-21 FFmpeg 下载界面