

第5章 视频处理技术及应用

学习目标

- (1) 掌握数字视频的概念和数字视频的特点。
- (2) 了解数字视频信息获取的基本原理和方法。
- (3) 掌握数字视频的编辑和处理。
- (4) 了解视频编辑软件的工作流程。
- (5) 掌握常见的视频编辑技巧。

视频就是一组随时间连续变化的图像。当连续图像的变化速率高于 24 幅画面/秒时,根据视觉暂留原理,这些画面将在人眼中产生平滑连续的视觉效果,这种连续变化的画面组则被称为视频或者动态图像等。

视频是通过摄像直接从现实世界中获取的,可帮助人们感性认识和理解多媒体信息所表达的含义。视频分模拟视频和数字视频。本章主要介绍数字视频的基本概念,视频采集、输出和压缩标准,视频处理软件 Premiere 的基本功能与应用方法。

5.1 视频处理技术概述

5.1.1 模拟视频与数字视频

视频(video)这个术语来源于拉丁语——“我能看见的”,通常指不同种类的活动画面,又称影片、录像、动态影像等,泛指将一系列静态图像以电信号方式加以捕捉、记录、存储、传送与重现的各种技术。按照视频的存储与处理方式不同,视频可分为模拟视频和数字视频两大类。

1. 模拟视频

模拟视频(analog video)属于传统的电视视频信号的范畴,其每一帧图像均是实时获取的自然景物的真实图像信号。模拟视频信号基于模拟技术以及图像显示的国际标准来产生视频画面,具有成本低和还原性好等优点。视频画面往往会给人一种身临其境的感觉,但缺点是不论被记录的图像信号有多好,经长时间的存储或多次复制之后,信号和画面的质量都会明显地降低。

电视信号是视频处理的重要信息源。电视信号的标准也称为电视制式。目前各国的电

视制式不尽相同,不同制式之间的主要区别在于刷新速度、颜色编码系统和传送频率等不同。目前,世界上最常用的模拟广播视频标准(制式)有中国、欧洲使用的 PAL 制,美国、日本使用的 NTSC 制及法国等国使用的 SECAM 制。

NTSC 标准是 1952 年美国国家电视标准委员会(National Television Standard Committee)制定的一项标准。其基本内容为:视频信号的帧由 525 条水平扫描线构成,水平扫描线每隔 1/30 秒在显像管表面刷新一次,采用隔行扫描方式,每一帧画面由两次扫描完成,每次扫描叫作一场,扫描一场需要 1/60 秒,两个场构成一帧。美国、加拿大、墨西哥、日本和其他许多国家都采用该标准。

PAL(Phase Alternate Lock)标准是联邦德国 1962 年制定的一种兼容电视制式。PAL 意指“相位逐行交变”,主要用于澳大利亚、南非、中国,以及欧洲大部分国家和南美洲。其屏幕分辨率增加到 625 条线,扫描速率降到 25 帧/秒,采用隔行扫描方式。

SECAM 标准是 Sequential Color and Memory 的缩写,该标准主要用于法国、苏联,以及和其他一些国家。SECAM 是一种 625 线、50Hz 的系统。

模拟视频信号主要包括亮度信号、色度信号、复合同步信号和伴音信号。PAL 彩色电视制式采用 YUV 模型来表示彩色图像,其中 Y 表示亮度,U、V 是构成彩色的两个分量,表示色差。与此类似,NTSC 彩色电视制式使用了 YIQ 模型,其中的 Y 表示亮度,I、Q 是两个彩色分量。YUV 表示法的重要性是其亮度信号(Y)和色度信号(U、V)相互独立,即 Y 信号分量构成的黑白灰度图与用 U、V 信号构成的另外两幅单色图相互独立。由于 Y、U、V 是独立的,因此可以对这些单色图分别进行编码。

2. 数字视频

数字视频(digital video)是相对于模拟信号而言的,指以数字形式记录的视频。数字视频有不同的产生方式、存储方式和播出方式。通过视频采集卡将模拟视频信号进行 A/D(模/数)转换——这个转换过程就是视频捕捉(或采集过程),再将转换后的信号采用数字压缩技术存入计算机磁盘中就成为了数字视频。

相对模拟视频而言,数字视频具有如下特点。

- (1) 数字视频可以不失真地进行无数次复制。
- (2) 数字视频便于长时间的存放而不会有任何的质量降低。
- (3) 可以对数字视频进行非线性编辑,并增加特技效果等。
- (4) 数字视频数据量大,在存储与传输的过程中必须进行压缩编码。

5.1.2 线性编辑与非线性编辑

1. 线性编辑

线性编辑是视频的传统编辑方式。视频信号顺序记录在磁带上,在进行视频编辑时,编辑人员通过放像机播放磁带并选择一段合适的素材,把它记录到录像机中的另一个磁带上,然后再按顺序寻找所需要的视频画面,接着进行记录工作,如此反复操作,直至把所有合适的素材按照节目要求全部记录下来。这种依顺序进行视频编辑的方式称为线性编辑。

2. 非线性编辑

非线性视频编辑是对数字视频文件的编辑,在计算机的软件编辑环境中进行视频后期编辑制作,能实现对原素材任意部分的随机存取、修改和处理。这种非顺序结构的编辑方式称为非线性编辑。

非线性编辑具有如下几个特点:

(1) 非线性编辑的素材以数字信号的形式存储到计算机硬盘中,可以随调随用,完成快速搜索,精确定位,图像的质量可以控制;

(2) 非线性编辑具备强大的编辑功能,一套完整的非线性编系统往往集成了录制、编辑、特技、字幕、动画等功能,这是线性编辑无法比拟的;

(3) 非线性编系统的投入相对较少,设备维护、维修及工作运行成本都较线性编辑大为降低。

非线性视频编辑的这些特点使得它已成为电视节目编辑的主要方式。

5.2 视频信号数字化

5.2.1 数字视频的采集

获取数字视频信息主要有两种方式:一种是利用数码摄像机直接获得无失真的数字视频;另一种是通过视频采集卡把模拟视频转换成数字视频,并按数字视频文件的格式保存下来。

一个数字视频采集系统由3部分组成:一台配置较高的多媒体计算机系统,一块视频采集卡和视频信号源,如图5-1所示。

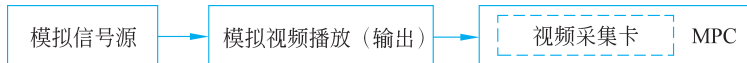


图 5-1 数字视频采集系统

1. 视频采集卡的功能

在计算机上通过视频采集卡可以接收来自视频输入端(录像机、摄像机和其他视频信号源)的模拟视频信号,对信号进行采集、量化,使之转换成数字信号,然后压缩编码成数字视频序列。大多数视频采集卡都具备硬件压缩功能。在采集视频信号时,首先在卡上对视频信号进行压缩,然后才通过 PCI 接口把压缩的视频数据传送到主机上。一般的视频采集卡采用帧内压缩算法把数字化的视频存储成 AVI 文件,高档一些的视频采集卡还能直接把采集到的数字视频数据实时压缩成 MPEG-1 格式的文件。

由于模拟视频输入端可以提供不间断的信息源,视频采集卡要采集模拟视频序列中的每帧图像,并在采集下一帧图像之前把这些数据传入计算机系统。因此,实现实时采集的关键是控制每一帧所需的处理时间。如果每帧视频图像的处理时间超过相邻两帧之间的相隔

时间,就会出现数据的丢失,即丢帧现象。采集卡先是把获取的视频序列进行压缩处理,然后再存入硬盘。也就是说,视频序列的获取和压缩是在一起完成的,免除了再次进行压缩处理的不便。

2. 视频采集卡的工作原理

视频采集卡的结构如图 5-2 所示。多通道的视频输入用来接收视频输入信号,视频源信号首先经 A/D(模拟/数字)转换器由模拟信号转换成数字信号,然后经视频采集控制器剪裁、改变比例后压缩存入帧存储器。输出模拟视频时,帧存储器的内容经 D/A(数字/模拟)转换器由数字信号转换成模拟信号输出到电视机或录像机中。

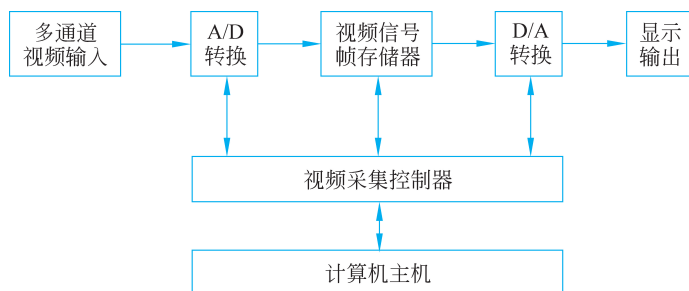


图 5-2 视频采集卡的结构

3. 视频采集卡与外部设备的连接

视频采集卡一般不配置电视天线接口和音频输入接口,不能用于直接采集电视射频信号,也不能直接采集模拟视频中的伴音信号。若要采集伴音,计算机需要装声卡。视频采集卡通过计算机的声卡获取数字化伴音,并同步伴音与采集的数字视频。

外部设备与视频采集卡的连接包括模拟设备视频输出端口与采集卡视频输入端口的连接,以及模拟设备的音频输出端口与多媒体计算机声卡的音频输入端口的连接。录像机(摄像机)用于提供模拟信号源,电视机用于监视录像机输出信号。完整的连接关系如图 5-3 所示。

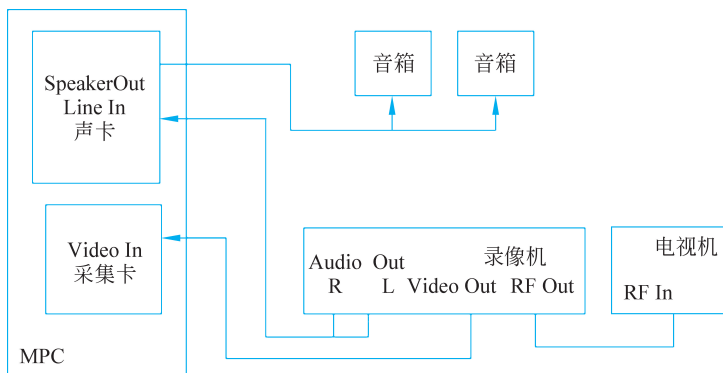


图 5-3 视频采集卡与外部设备的连接

如果 VHS 录像机具有 Video Out、Audio Out(R、L)和 RF Out 输出端口,则采用以下

连接:录像机的 Video Out 与采集卡的 Video In 相连;录像机的 Audio Out 与声卡的 Line In 相连;录像机的 RF Out 与电视机的 RF In 相连;声卡的 Speaker Out 与音箱相连。按照这种方式连接好之后,如果软件设置正确,那么通过多媒体计算机的音箱可以监视采集的伴音情况,而采集的视像序列则可以直接显示在多媒体计算机的显示器上。

5.2.2 数字视频的输出

数字视频的输出是数字视频采集的逆过程,即把数字视频文件转换成模拟视频信号输出到电视机上进行显示,或输出到录像机记录到磁带上。与视频采集类似,这需要专门的设备对数字视频进行解压缩及 D/A 变换,完成数字数据到模拟信号之间的转换。根据不同的应用和需要,这种转换设备也有多种。将集模拟视频采集与输出于一体的高档视频采集卡插在 PC 机的扩充槽中,连上较专业的录像机,则可提供高质量的模拟视频信号采集和输出,用于专业级的视频采集、编辑及输出。

另外,还有一种称为 TV 编码器(TV Coder)的设备,它的功能是把计算机显示器上显示的所有内容转换为模拟视频信号并输出到电视机或录像机上。这种设备的功能较为简单,适合于普通的多媒体应用。

5.3 数字视频压缩标准与文件格式

5.3.1 数字视频数据压缩标准

数字视频数据量通常非常庞大,为了方便存储和传输,常需要采用有效的途径对其进行压缩。基于视频数据的冗余性,人们分析研究出一系列编码压缩算法,这些方法大致可分为帧内压缩和帧间压缩两种。

(1) 帧内压缩。该压缩在压缩图像帧时,仅考虑本帧的数据而不考虑相邻帧之间的冗余信息,帧内一般采用有损压缩算法,达不到很高的压缩比。

(2) 帧间压缩。该压缩基于视频或动画(连续的)前后两帧的极大相关性(即连续的视频其相邻帧之间具有冗余信息)特点来实现。通过压缩时间轴上不同帧之间的数据,进一步提高压缩比。

与音频压缩编码类似,为了使图像信息系统及设备具有普遍的互操作性,一些相关的国际化组织先后审议并制定了一系列有关图像编码的标准。其中,MPEG 系列标准由运动图像专家组(Moving Picture Experts Group)制定。

MPEG 系列标准包含 MPEG-1、MPEG-2、MPEG-4、MPEG-7 和 MPEG-21 等 5 个具体标准,每种编码都有各自的目标问题和特点。

MPEG-1 标准的目标是以约 1.5Mb/s 的速率传输电视质量的视频信号,亮度信号的分辨率为 360×240 ,色度信号的分辨率为 180×120 ,帧速为 30 帧/秒。这是世界上第一个用于运动图像及其伴音的编码标准,主要应用于 VCD,其音频第 3 层即 MP3 广泛流行。该标

准于1988年5月提出,1992年11月形成国际标准。

MPEG-2标准于1990年6月提出,1994年11月形成国际标准。该标准的视频分量的位速率范围为2~15Mb/s,分辨率有低(350×288)、中(720×480)、次高(1440×1080)、高(1920×1080)等不同档次,压缩编码方法也分为从简单到复杂等多个等级,广泛应用于数字机顶盒、DVD和数字电视。

MPEG-4标准于1993年7月提出,1999年5月形成国际标准。该标准是一种基于对象的视/音频编码标准。采用MPEG-4技术,一个场景可以实现多个视角、多个层次、多个音轨,以及立体声和3D视角,这些特性使得虚拟现实成为可能。MPEG-4标准制定了大范围的级别和框架,可广泛应用于各行各业。

MPEG-7标准于1997年7月提出,2001年9月形成国际标准。该标准是一种多媒体内容描述标准,定义了描述符、描述语言和描述方案,支持对多媒体资源的组织管理、搜索、过滤和检索等,便于用户对其感兴趣的多媒体素材内容进行快速有效的检索,可应用于数字图书馆、各种多媒体目录业务、广播媒体的选择和多媒体编辑等领域。

MPEG-21标准几乎与MPEG-7标准同步制定,于2001年12月完成。MPEG-21标准的重点是建立统一的多媒体框架,为从多媒体内容发布到消费所涉及的所有标准提供基础体系,支持连接全球网络的各种设备透明地访问各种多媒体资源。

H.264是一种视频高压压缩技术,全称是MPEG-4 AVC,即活动图像专家组-4的高等视频编码。它是由国际电信标准化部门(ITU-T)和规定MPEG的国际标准化组织(ISO)及国际电工协会(IEC)共同制定的一种活动图像编码方式的国际标准格式。H.264的优势主要在于超高压压缩率、高国际标准和公正的无差别许可制度。

5.3.2 数字视频文件格式

数字视频文件格式大致可分为普通视频文件格式和网络流式视频文件格式两类。

1. 普通视频文件格式

1) AVI格式

AVI(Audio Video Interleaved)是一种音视频交叉记录的数字视频文件格式,运动图像和伴音数据以交替的方式存储。这种音频和视像的交织组织方式与传统的电影相似,包含图像信息的帧顺序显示,同时伴音声道也同步播放。

AVI文件结构不仅解决了音频和视频的同步问题,而且具有通用和开放的特点。它可以在任何Windows环境下工作,而且还具有扩展环境的功能。用户可以开发自己的AVI视频文件,在Windows环境下随时调用。

AVI一般采用帧内有损压缩。可以用一般的视频编辑软件(如Adobe Premiere)进行再编辑和处理。这种文件格式的优点是图像质量好,可以跨平台使用;缺点是文件体积较大。

2) MPEG格式

MPEG(Moving Picture Expert Group)/MPG/DAT的具体格式后缀是mpeg、mpg或dat,家用VCD、SVCD和DVD使用的就是MPEG格式文件。

将MPEG算法用于压缩全运动视频图像,就可以生成全屏幕活动视频标准文件——

MPG 文件。MPG 格式文件在 1024×786 的分辨率下可以用 25 帧/秒(或 30 帧/秒)的速率同步播放全运动视频图像和 CD 音乐伴音,并且其文件大小仅为 AVI 文件的六分之一。MPEG-2 压缩技术采用可变速率(VBR-Variable Bit Rate)技术,能够根据动态画面的复杂程度,适时改变数据传输率获得较好的编码效果。DVD 就是采用了这种技术。

MPEG 的平均压缩比为 $50:1$,最高可达 $200:1$ 。同时,其图像和音响的质量也非常好。MPEG 标准包括 MPEG 视频、MPEG 音频和 MPEG 系统(视频、音频同步)3 个部分,MP3 音频文件就是 MPEG 音频的一个典型应用,而 VCD、SVCD、DVD 则是全面采用 MPEG 技术所产生出来的新型消费类电子产品。

3) MOV 格式

MOV(Movie digital video technology)是美国 Apple 公司开发的一种视频文件格式,默认的播放器是 QuickTime Player,具有较高的压缩比和清晰度,并且可跨平台使用。

2. 网络视频文件格式

1) RM 格式

RM 是 Real Networks 公司开发的一种流媒体文件格式,也是目前主流的网络视频文件格式。Real Networks 制定的音频、视频压缩规范称为 Real Media,相应的播放器为 Real Player。

2) ASF 格式

ASF(Advanced Streaming Format)格式是微软公司前期的流媒体格式,采用 MPEG-4 压缩算法,这是一种可以在因特网上实时观看的视频文件格式。

3) WMV 格式

WMV(Windows Media Video)格式是微软公司推出的采用独立编码方式的视频文件格式,也是目前应用最广泛的流媒体视频格式之一。

5.4 视频编辑软件 Premiere

在一个完整的非线性编辑系统中,硬件只能提供对音频、视频数据的输入、输出、压缩、解压缩、存储等工作的处理环境,视频、音频的编辑则要通过非线性编辑应用软件才能实现,即数字视频的后期编辑工作主要依靠视频编辑软件来完成。

目前市场上的非线性编辑软件种类较多,比较流行的是 Adobe 公司的 Premiere 系列和 Ulead 公司的 Video Studio(会声会影)系列,它们可以和大多数的视频采集卡配合使用,其工作原理基本相似,都采用了时间轴和各种素材轨的编辑方法。本节主要介绍 Adobe Premiere 数字视频编辑软件的功能和使用。Adobe Premiere 软件为剪辑人员提供了非常实用的工具,能让用户得心应手地完成剪辑任务。

5.4.1 Premiere 工作界面

Premiere 的界面如图 5-4 所示。



图 5-4 Premiere 界面

项目面板的主要功能是对素材进行存放和管理。编辑视音频素材时,首先要导入这些素材并进行相应的设置与管理,以便分类和安排编辑次序。项目管理还提供了新建文件夹、创建素材、搜索等功能,方便用户维护和使用素材。

监视器视窗用于实时预览影片和剪辑影片。默认的监视器采用双显示模式,即包含源监视器和节目监视器。素材监视器负责存放和显示待编辑的素材,节目监视器实时预览已经编辑完成的影片。

时间线面板是 Premiere 中核心界面之一。在时间线面板中,可根据脚本将视频、音频、图像等有组织地剪辑在一起,并加入转场、特效、字幕,从而制作出精美的影片。

特效面板为用户提供了丰富的视/音频转场和特效,极大地丰富了画面语言的处理,为影片的创作提供了更为广阔的空间。该部分由预置、音频特效、音频过渡、视频特效、视频过渡等几大块内容组成。后面的实例学习将充分展示这些特效带来的丰富体验。

工具面板提供了影片剪辑和动画关键帧设置所需要的一些工具,包括选择工具、涟漪编辑工具、滚动编辑工具、速度调整工具、剃刀工具、滑动工具、钢笔工具等。

5.4.2 项目管理

在整个 Premiere 非线性编辑的工作流程中,前期的项目设置及管理非常重要。

1. 创建项目

启动 Premiere CS 程序时,系统会询问是否新建或者打开项目,如图 5-5 所示。选择新建项目后将打开“新建项目”对话框,如图 5-6 所示。单击“确定”按钮即可进入空白的工程

界面。此时,执行“文件”→“新建”→“序列”命令或者按 Ctrl+N 键打开如图 5-7 所示的对话框。Premiere 给出了多种预置的视频和音频配置,有 PAL 制、NTSC 制、24P 的 DV 格式、HDTV 等。可以在序列预设选项卡中选择 DV-PAL 制式标准 48kHz,也可以选择常规面板进行自定义,如图 5-7 所示。

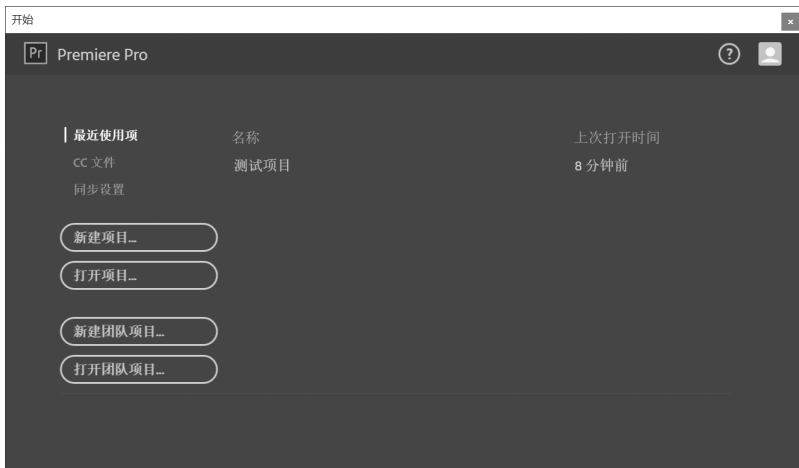


图 5-5 启动界面

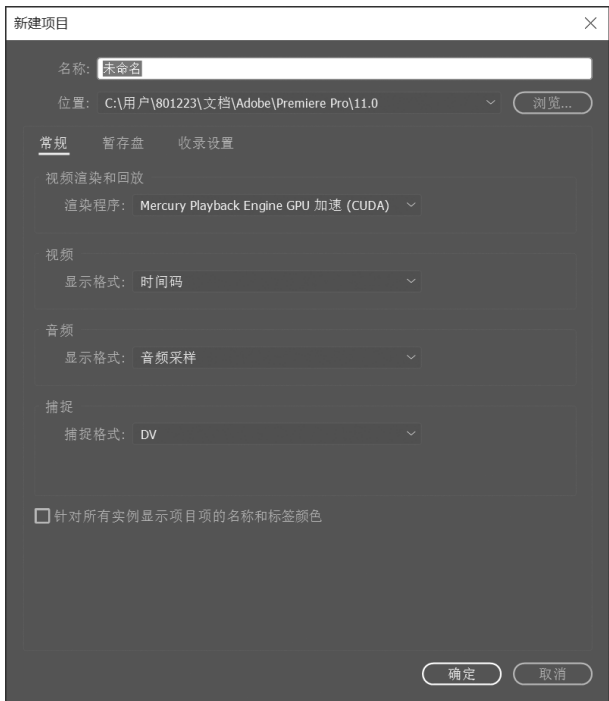


图 5-6 新建项目设置

每种预设方案都包括文件的压缩类型、视频尺寸、播放速度、音频模式等方面的信息,还可以在“设置”面板中进行自定义设置,如图 5-8 所示。下面对该面板做简要介绍。

(1) 编辑模式。在序列预设选项卡中选择某种预设后,其编辑模式会同时显示出来,可以根据需要改变编辑模式。

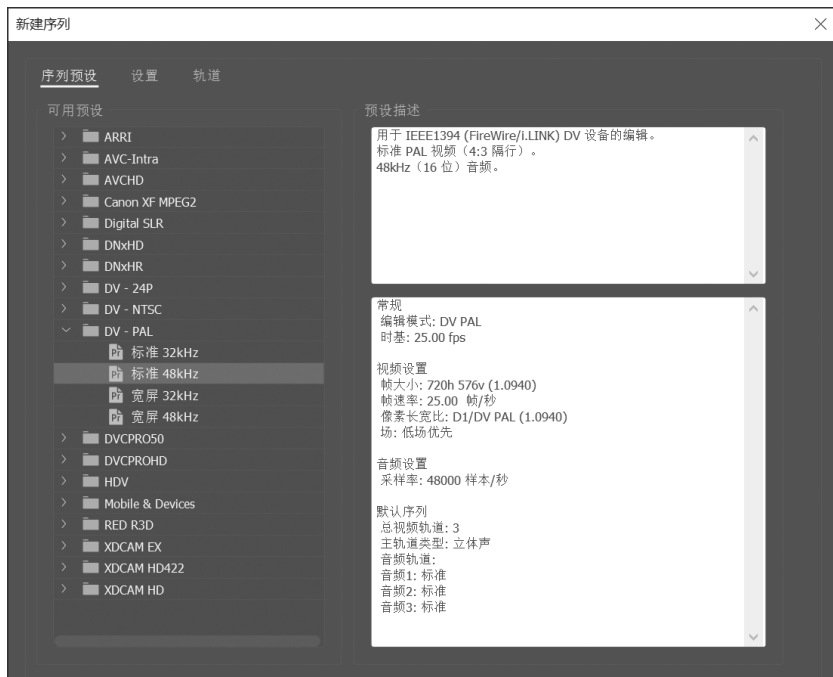


图 5-7 预设方案

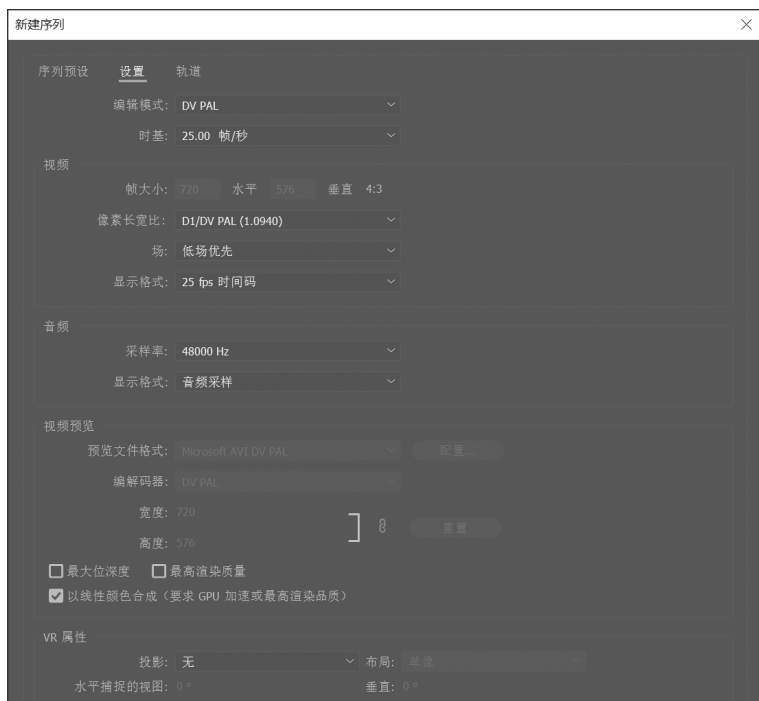


图 5-8 常规选项

(2) 时基。即每秒播放多少帧,一般 PAL 制式 25 帧,NTSC 制式 30 帧。我国采用的是 PAL 制,因此,一般的电视节目都是 25 帧。