

MPEG-2 TS 码流编辑的原理与应用

在当今数字媒体不断发展、新媒体业务不断涌现的前提下，实践证明襁褓中的新媒体只有两种经营方略可供选择：或是购买并集成整套节目，或是低成本深加工新节目，再不可能去按照传统生产模式去自采自编。低成本的节目生产制作与发布，不仅成为数字媒体经营的主要手段，也成为传统媒体“改革工作流程”的重要举措，进而促成了对新型工作母机和简捷快速流程的迫切需求。

在辽宁新媒体多业务综合服务平台上，先于国际和国内应用了 MPEG-2 传输流快速剪辑编辑系统（以下简称码流快编）。这项由辽宁电视台与深圳奥维迅公司在 2003 年 10 月联合开发的新技术，为数字媒体低成本节目的制作、推广和运营提供了高效生产工作母机。尽管担负此项目源代码开发的奥维迅公司出现了经营问题，在技术推广的中间环节发生梗塞，但并不能说明此项技术走到了尽头。回顾 3 年的应用实践及研发成果，需要的不是扬弃，而是演进的升级，否则就是对可调控资源的莫大浪费。特别是针对第二代信源编解码国标 AVS-P2 的更新换代，很可能成为多业务内容整合的新一代产品的突破口。

一 工作原理

1. 功能目标

码流快编的应用目标是，通过对开放视频的采集，将 DVB-S 或 C 的传输流（Transport Stream, TS）节目作为信源，直接进行剪辑处理，再经过人工创意后，整合为新主题内容的新节目，以便直接进入频道集成或编辑频道节目播出，快速实现数据层的内容整合，不仅简捷了采集制作的工作流程，而且为丰富媒体内容资产开辟了一条捷径。因为码流快编的工作流程无需先以解码后的视频记录于磁带，再以磁带上载编辑机，经编辑后再下载成为磁带，再编码复用成为新内容的新节目。即便数字化完成以后，视频数据流仍不能用于经复用的数字传输，还需编码、转码、打包等传输格式化以后，才能在数字信道上传输。而采用码流快编以后，不仅避免了解码后再采样编码所形成的视频损耗，还避免了在 1:1 时间的上下载中所造成的效率损耗。更重要的是在视频内容整合中，一次性完成音/视频同步剪切、字幕处理和音/视频数据打包复用等连续作业。所以，它能够提高生产效率 60% 以上。必要时还可进行节目包装的特技编辑，直接创建数据级和文件级的互联互通内容交换平台，在媒体资产管理下，顺利实现网络化与智能化的节目配送与发布。

由于码流快编是针对以 TS 为信源的再编辑系统，所以实行“高来高走，低来低走”，或是“高来低走”的应用策略，即高码率对应高码率（包括兼容高清），低码率对应低码率，但码率连续可调，以适应高码率对应低码率的应用。理论和应用都说明，对比源节目和成品节目，经剪切和编辑处理的图像保持了同等的视频质量，成为不劣化图像的创新工作流程和新型工作母机。

2. 设计特征

码流快编的低成本与高效率来自它的主要特征。常规的非线性编辑机是在编码一侧做文章，通过采集编码卡实现视频图像多层多轨的非线性编辑；而码流快编则是在解码一侧下功夫，通过对 TS 拆包还原为基本码流（Elementary Stream，ES），而后经编辑再封包成为 TS，快速实现视频内容整合的业务应用。因而，码流快编除了运用非线性图像处理技术外，最大的特点是运用了 MPEG-2 和 DVB 的系统原理，在此基础上进行图像处理的应用开发。

通过图 1 的系统概念，可以看出码流快编运用了一个逆向思维的方法，利用 DVB 系统传输的发/收互逆关系，将收端的单节目传输流（SPTS）作为信源，通过 ES 实现以视频非线性处理的目标功能。这时的解决方案有两个：一是在 TS 基础上直接进行图像处理的基础开发，实现与常规非编一样的操作界面和编辑习惯，这样虽好，但是没有参照模型以及可利用的技术资源，必须从零开始的重写源代码；二是将 TS 转码为 ES，以 ES 帧结构和句法格式还原成为符合 ITU-R.601 建议的原始视频帧，就可以很方便地利用或附加现有的常规非编技术，以无卡站形式实现编辑。码流快编同时采用了两种方案，一方面独立开发新产品系列，以适应于快速发展的数字电视业务需求，另一方面提高系统的兼容性，向后兼容传统非编，有利于在媒体资产管理下实现互联互通和投资保护。图 2 说明了方案二的 ES 还原并显示原始视频的 GOP 帧结构。

通过图 2 可以认为，既然能够形成 IBBP 的句法帧，再转换成为全 I 帧格式并不难。这样一来，就可以利用原有的非线性编辑技术，连续处理长与短 GOP、全 I 帧与 IBP 帧结构、可变与固定码流的节目素材，实现不同节目格式的快编与混编。由此可见，码流快编的开发技术难点，是结合 DVB 系统与 MPEG-2 标准，从译码过程中读出元数据，实现基于解码器的图像处理技术，并能兼容以编码卡为基础的非线性编辑技术。因而被业内称为“第一个吃螃蟹者”。

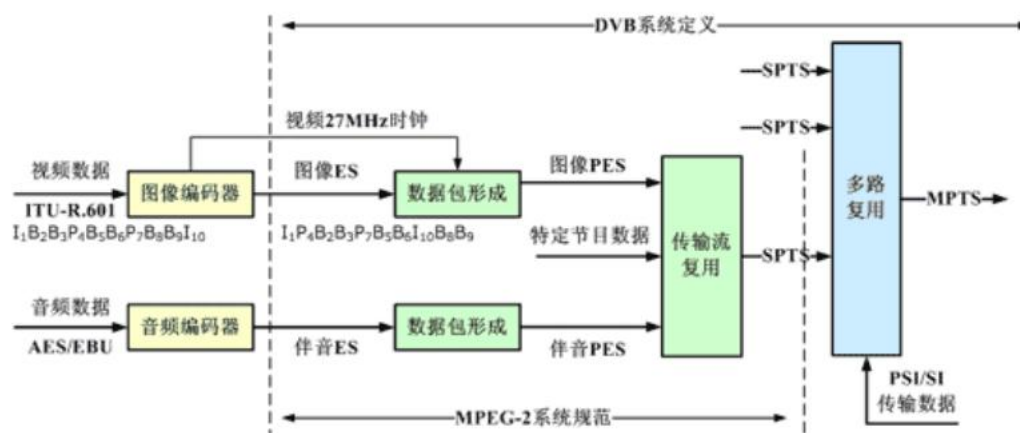


图1 DVB和MPEG-2系统中数据流的区分

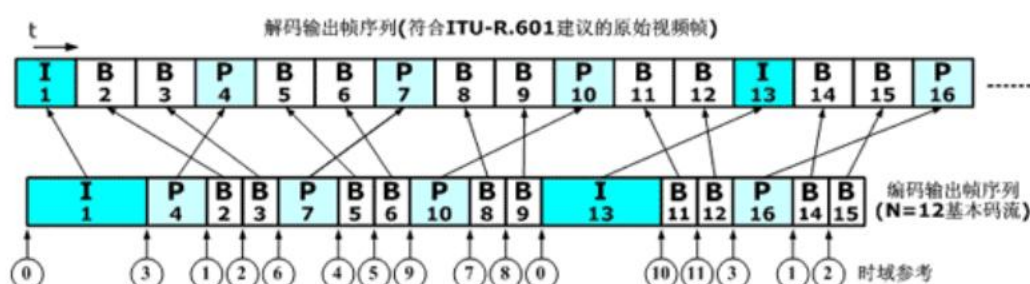


图2 ES句法和帧结构的还原

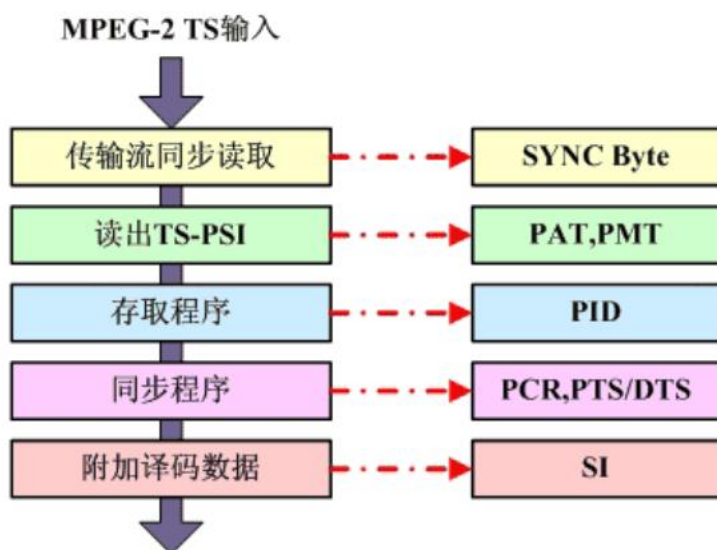


图3 解码流程



3. 数据变换

既然要把 TS 作为节目源进行编辑,就需要将 TS 包中的数据变换为非线性编辑所能使用的元数据和视频流。它们是以码流快编作为工作母机进行生产的真正原料。

数据变换的第一个过程是拆 DVB 复用包。如图 3 所示, DVB 的解码流程中分层译出了许多数据信息,如同步字节 (Sync Byte)、节目特定信息 (Program Specific Information, PSI) 中 PAT、PMT、NIT 等列表、包识别 (Packet Identification, PID)、节目时钟基准 (Program Clock Reference, PCR) 及 PTS/DTS (后详解) 和业务信息 (Service Information, SI) 等部分。这些信息不仅与 DVB PSI/SI 直接相关,与 MPEG-2 的句法结构也直接相关,它规范地传递了再生码流中音/视频所需的 MPEG-2 列表数据,通过这些信息的引导,准确进入 MPEG-2 系统层的进一步译码。

第二个过程是拆 MPEG-2 系统复用包。MPEG-2 系统定义了一个的码流层次化结构句法规则,以便于误码处理、随机搜索以及内容编辑。它自上而下依次分为图像序列层 (Video Sequence Layer, VSL)、图像组层 (Group of Pictures Layer, GOP/L)、图像层 (Picture Layer, PL)、像素条层 (Slice Layer, SL)、宏块层 (Macro block Layer, ML)、像块层 (Block Layer, BL) 等 6 个层次,分别赋予每个层次不同的功能。图 4 说明了 MPEG-2 体系的句法结构,通过这种分层排列的结构特征, MPEG-2 提高了系统的灵活性和管理效率,使得每一层都可以用来支持一个特定的功能。码流快编大部分应用于 VSL、GOP/L、PL 层,特别是在 GOP/L,需要在还原时规范翻译 PID、PCR、PTS/DTS 等信息,将同步信息、闭合标记、断链标记等数据准确插入 GOP 图像组,以形成每组 GOP 数据流的起点,才能保证图像帧的精确和连续帧的同步接续。

第三个过程是将拆包后所得数据信息,以规范的装填还原 MPEG-2 ES 以及元数据,因而装填数据是码流快编的重要技术环节。

(1) 装填复用的基本码流包

依据 MPEG-2 TS 规范结构 (如图 5 所示),复用的基本码流包 (Packetized Elementary Stream, PES) 是由包头、自适应区的 ES 特有信息和包数据 3 部分所组成。由于包头和 ES 特有信息二者可合成 1 个数据头,因而可认为 1 个 PES 包是由包头和包数据 (有效载荷) 2 个部分所组成。对有线、卫星、地面广播网接收的 TS,经过解调和解扰处理后选取包长为 188B 的 SPTS,并从包头中提取相关信息,以 PID 区别不同 SPTS 包,以连续计数器的顺序计对标注 PID 的 TS 包重建一个独立分组的 PES。根据自适应区中的填充数据,装填到不为 TS 包整数倍的 PES 包中,以保证 PES 变长包的完整性。同时,依据包头及自适应区内的同步字节、原始程序参考时钟 (Original Program Clock Reference, OPCR)、PCR 等同步和识别信息,提供 27MHz 的解码同步时钟,装填共同时间基准、独立时间基准、可变包长和有效载荷等数据。

装填成为 PES 对码流快编具有格外重要的意义,因为 PES 包内含音/视频的 ES 以及包括 PID 的 12 个包头识别标志,当对 PES 包的起点不能精确定位、对 PES 包头标志符不

能准确识别时，就不能保证拆包后再打包的图像无缝接续和声画同步。这是在初期研发过程中遇到并获得突破 的技术难点。

（2）装填基本码流

根据 MPEG-2 规范的 PES 包结构（如图 6），还需进一步装填为连续 ES 流。ES 是指只包含 1 个信源的数据流，即视频数据流或 音频数据流。每个 ES 由若干个缓存器的特定存取单元（Access Unit，AU）所组成，而每个视频 AU 或音频 AU 都是由头部和编码数据的 2 部分。1 个 AU 相当于编码的 1 帧视频图像或 1 个音频帧的取样。PES 的包头为 恢复 ES 提供了向导。

对于 PES 包头，具有 ES 特有信息的显示时间标记（Presentation Time Stamp，PTS）、解码时间标记（Decode Time Stamp，DTS）标志、基本流时钟基准（Elementary Stream Clock Reference，ESCR）信息标志、基本流速率信息标志、数字存储媒体（Digital Storage Media，DSM）的特技信息标志等等，其中，唯有 PTS/DTS 标志，是解决视音频同步显示，防止输入缓存器上溢或下溢的关键所在。

在装填过程中，数据定位指示符引导 PES 还原所包含的视频、音频及所属其它数据流（如同步、数据和数据通道等），包头识别标志 的 PTS/DTS 指示了可变长度包数据的帧同步时间，当区分音/视频和其它数据以后，依据 PTS/DTS 对视频帧分配给特定的 AU。其中尤以独立时间基准 是还原 ES 的同步基础。对于 PES 包数据，一方面通过扩展标志的数据包计数器，引导恢复数据流，另一方面利用循环冗余校验（Cyclic Redundancy Check，CRC）辅助检测并纠正可能存在的数据包丢失。

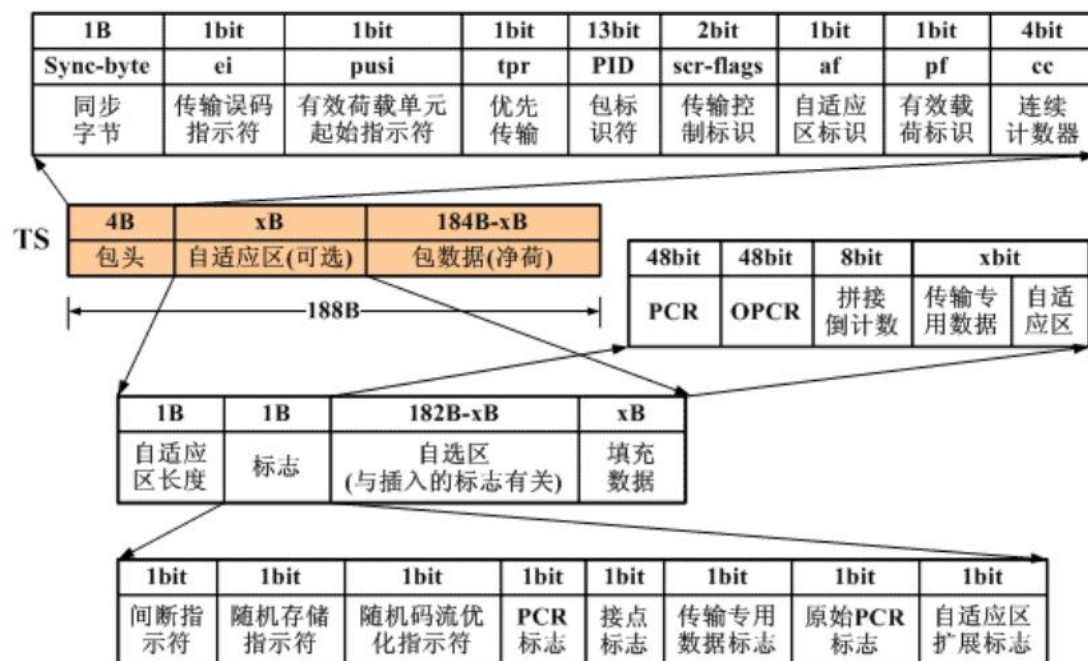


图5 TS结构示意图

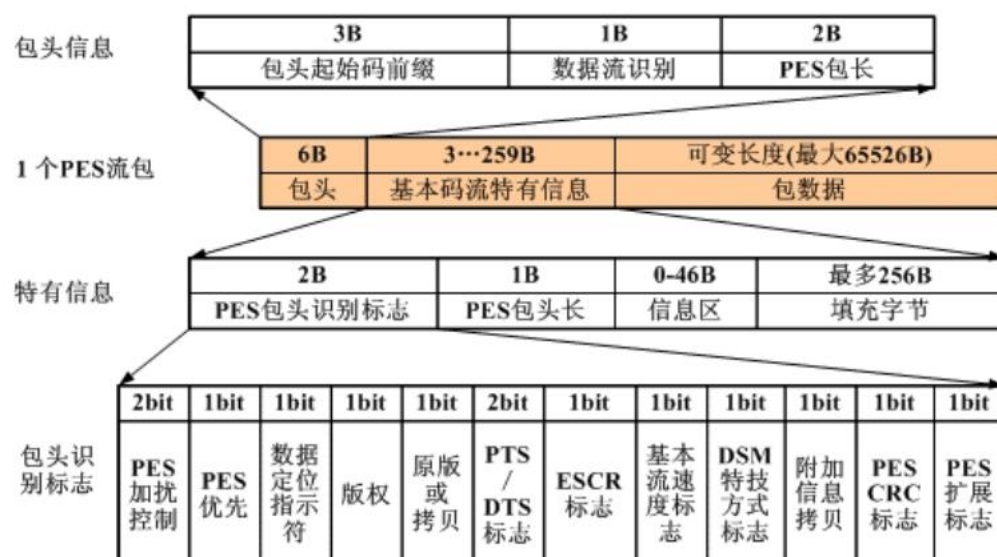


图6 PES包结构

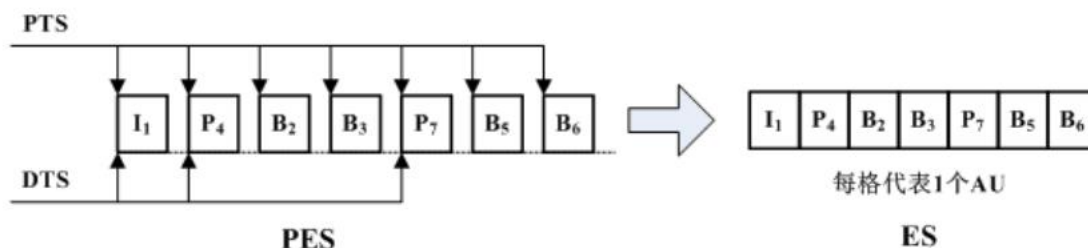


图7 PES解包为ES示意图

4. 精确帧定位

帧定位是精确编辑的基础，而精确的帧定位来自于精确的帧同步。在 ES 上实现逐帧精确的编辑，首要问题是实现精确帧的同步。

如前所述，装填后的 ES 变成仅含有 1 种性质的 PES 包，或视频 ES，或音频 ES。以视频为例，图 7 表明了 PTS/DTS 标示的独立同步时间基准指示下，顺序装填再顺序读出，形成连续 ES 的 I1P4B2B3P7B5B6 (N=7) GOP 组帧顺序。由于 PES 的数据分组是可变长度的数据包，但它的最大包数据容量为 65526Byte。因此，在码流快编中必须为每个 AU 准备必不可少的缓存空间。

如图 7 所示，PTS 表明图像帧出现在目标解码器 (System Target Decoder, STD) 的时间，DTS 表明将存取单元全部字节从 STD 的 ES 解码缓存器移出的时刻。当以 PTS/DTS 为独立时间基准，定位和标志 PES 的 AU 起始点后，对 UA 依次组成图像帧序为 I1P4B2B3P7B5B6 I10B8B9 的 ES。对于 I、P 帧而言，PES 的图像帧序为 I1P4B2B3P7B5B6I10B8B9，应该 P4 比 B2、B3 在先，但显示时 P4 一定要比 B2、B3 在后，这就必须重新排序。在 PTS/DTS 时间标志指引下，将 P4 提前插入数据流，经过缓存器重新排序，重建视频帧序 I1B2B3P4B5B6P7B8B9I10。显然，PTS/DTS 是表明确定事件或确定信息，并以专用时标形态确定事件或信息的开始时刻。

值得注意的是，虽然在 PES 中应该每个 I、P、B 帧的包头都具有一个 PTS 和 DTS，但由于 B 帧的解码时间和显示时间存在一致性，因而对 B 帧而言，PTS 与 DTS 具有相同作用，无须 DTS，只须 PTS。音频数据包也很有特点，虽然它可以含有多个存取单元，但由于它必须按照时间顺序传送，所以音频包头中也不含 DTS，只有 1 个 PTS。

由此可见，PTS/DTS 不仅直接关联帧定位，而且直接关系到码流快编的帧精确编辑。为此，码流快编的帧定位着重处理以下几个帧序特征：

- * 当处理某个含有 1 个 I 帧的存取单元时，在其包头文件中应有 DTS 和 PTS，且 2 个标记之间的时间间隔为 1 个图像周期，那么在双向编码时的 P 帧应在 I 帧之后，在包头文件中就应有 1 个 DTS 和 1 个 PTS，而这 2 个标记之间应存在 3 个图像周期的间隔，这样才能插入 2 个 B 帧。

- * 当前处理 IPBB 时，I 帧应延迟 1 个图像周期，P 帧应延迟 3 个图像周期，而 2 个 B 帧则无需延迟，于是，最后的显示顺序就成为了 IBBP 帧序。

- * 若需要改变 GOP 结构时，例如在 I 与 P 帧之间需要存在更多的 B 帧，则需调整 B 帧 DTS 与 PTS 之间的时间间隔。

- * 当处理 IPBB 序列后，必在第 1 个 B 帧之前先对 I 和 P 帧进行解码，但每次只能对一帧图像进行解码，因此需要先解 I 帧并暂存后，待 P 帧被解码时，随即读出 I 帧，而后再读 B 帧。

* 在处理包头时，对 PTS/DTS 指示的某图像帧，特殊关注是否只有 1 个 PTS 时间标记，是否含有 PTS 和 DTS 的 2 个时间标记，以便快速确认 I、P 与 B 帧，以及音频包。

当作为编辑节目源的 MPEG-2 TS 被拆包以后，以 PTS/DTS 实现视频帧的精确定位，就能以帧精确的剪切编辑实现码流快编功能，同时，也能以 I1P4B2B3P7B5B6I10B8B9 帧序显示 I1B2B3P4B5B6P7B8B9I10 帧序。那么，在此基础上嫁接现成的 MPEG-2 IBP 非线性编辑的成熟技术，既不需要高价的采集编码卡，也不需要上下载的转码，从采集、转码、拆包剪切，到特技处理、打包上传，一气呵成地构成了码流快编系统。



图9 上载工作站与模拟磁带接口的操作界面



图10 目标文件的合成属性选项



图11 码流快编的非线性编辑时间轨

二 实际应用

考虑到码流快编的兼容性和通用性，以及互联互通网络平台内容交换的需要，系统是在 Windows 平台上开发应用，系统流程详见图 8。

事实上就应用层而言，看不出码流快编与常规非编究竟有什么区别。只是对照图 8，才能看出常规应用时是以传统磁带记录的内容为节目源，这就需要围绕采集编码卡，设立上下载工作站，通过将模拟节目采集编码为数据流以后，再进行剪切等编辑。因而，目前还需对广泛应用的模拟视频提供接口选项，附加上下载工作站。它的应用界面及其采集参数选项参见图 9。图 9、10 表示数据输入的采集操作界面，显示目标文件格式初始化选项与合成属性，以维系连续应用的工作流程。

经采集所合成的码流格式文件，可以同时转换成为可供交换的各种流格式和连续可调的速率，以适应各种内容编辑的整合策略需求，并可利用网络，对连续生产的其它工作站输送和交换素材，直至包括特技的特殊内容加工。这些交换文件的合成属性以及应用格式选项参见图 10。

虽然在码流快编上进行快速编辑并合成文件时，与常规操作不无不同，但是通过对 TS 文件的快速正、反向搜索，以每一帧的入点、出点的精确定位，进行非线性编辑的剪切操作，制作成为新内容体裁的新文件，却被广电总局主管节目的领导所认可，被称之为“海量空中节目的过滤器”。同时，它还可以根据内容的需要，通过内置的迭加台标、栏目、字幕等字幕软件，直接参与编辑节目内容信息编辑或局部遮盖，并可联立采集工作站，形成生产流水线，进行不同任务种类的专项加工。如图 11 所示，在时间轨上的不同视频、音频、字幕等，均可按帧精确的快速合成为新节目或新素材。

如图 12 所示，当合成新节目或素材时，文件格式的也可以根据使用的需要加以选择。如果作为 DVB 播出节目的成品，则可选择“MPEG-2TS 流（DVB 格式）”选项；如果需要作为进一步深加工的素材，则可选择“MPEG-2ES 流（M2V+MPA）”选项，以素材集的形式存放于共享应用的集中存储。这样编辑后的节目素材或成品，理论上可以保证与节目源完全相同的视频质量。

还值得一提的是，只要在此基础上再增加一块廉价的 Matrox-RGX100 采集卡，就可以进一步集成无限轨、无限层的第二代非线性编辑技术，成为多格式混编、混排的超级深度特技编辑机，为动漫、游戏等新媒体节目的快速加工创造工具。

三 结语

无论码流快编对节目制作的现实意义存在何种褒贬，但在一次数字版权还没有进阶到二次版权法规的时候，它的存在具有特殊重要意义。毕竟能够通过仅 12 名（前期 6 人）的人员投入，以低成本、高效率的节目加工，为辽宁电视台提供了 18 套全省的数字电视节目，而且维系了近 3 年的节目生产，至今赢得了全省近百万数字电视用户的向往与肯定。试想一下，倘若按照常规作法，只播出 1 套节目用 12 个编辑就很可能要“累死人了”。

实践证明，当国外的政治体制和法规观念与国内存在现实差异的时候，国际上不屑一顾的产品并不是国内市场没有需求；当国际与国内 市场趋于接轨的时候，对瞬息万变的市场不屑一顾也不是新技术的初衷；当码流快编因某种原因并没有得到推广的时候，不屑一顾地扬弃它的内核就不是实事求是。 最恰当的对策是继承技术原理的设计理念，与时俱进地升级新产品，才能获得广电和企业的整体效益。