DIGITAL PROGRAM INSERTION FOR LOCAL ADVERTISING

Mukta Kar, Ph.D., Majid Chelehmal, Ph.D., Richard S. Prodan, Ph.D.
Cable Television Laboratories

Abstract

Current advertising insertion systems enable cable headends and broadcast affiliates to insert locally-generated commercials and short programs in a remotely distributed regional program before delivery to the home viewer. The revenue generated by the local ads and short programs is very significant. Current ad insertion systems are analog in nature, switch between uncompressed video sources, and use digital video compression for local storage only.

Digital compression provides digital audio, video, and data with superior quality and efficiency to existing analog means. The application considered in this paper involves insertion of locally-generated compressed digital commercials and short programs into a digital channel containing previously multiplexed and digitally compressed programs. However, combining compressed video streams presents additional challenges for seamless insertion of ancillary programs. It is difficult to splice a compressed digital bitstream compliant to the MPEG-2 standard without adversely affecting the resultant display due to decoding failures at compressed bitstream discontinuities.

This paper presents a brief overview of existing analog and digital program insertion systems.Various solutions, including limitations not present in current analog systems, are discussed. Flexibility, implementation complexity, and network operational constraints are also discussed for various potential digital program insertion methods.

INTRODUCTION

Cable advertising generates significant revenue for the cable industry. National ad-sales (advertising sales) and local ad-sales in early 1980 were $50 million and $8 million, respectively. National ad-sales for 1997 are expected to top $5 billion. Revenues from local ad-sales are expected to top $2 billion. These figures underscore the growing importance of advertising for Multiple Systems Operators (MSOs).

Current advertisement (ad) or commercial insertion methods are essentially analog, using strictly uncompressed video. A hybrid system stores local spots in compressed form, but decodes upon playback into uncompressed video prior to insertion. Analog insertion stores the commercial on tape in analog format. Hybrid insertion stores commercials in digital and compressed format on a server and converts them back to analog format before insertion into analog channels.

The advent of television signal compression technology and its standardization (MPEG-2), coupled with digital modulation, have ushered in a new era in the television broadcasting industry. One of the most important benefits is the bandwidth efficiency in spectrum utilization compared to its analog counterpart. An existing 6 MHz analog channel can be used to send multiple digital channels with equal or better picture and sound quality. This indicates that the existing limited number of analog channels can be transformed into a larger number of viewing or logical channels. These logical channels can be used for audio, video, and data services. To take advantage of this benefit, MSOs are gradually moving to a digital tier and adding more programs.

Cable operators also expect that their ad sales revenue will go up by inserting additional commercials targeted to a more specific audience or market segment in the added digital channels. Unfortunately, the devices used for insertion of commercials in analog channels will not work in compressed digital channels without significant

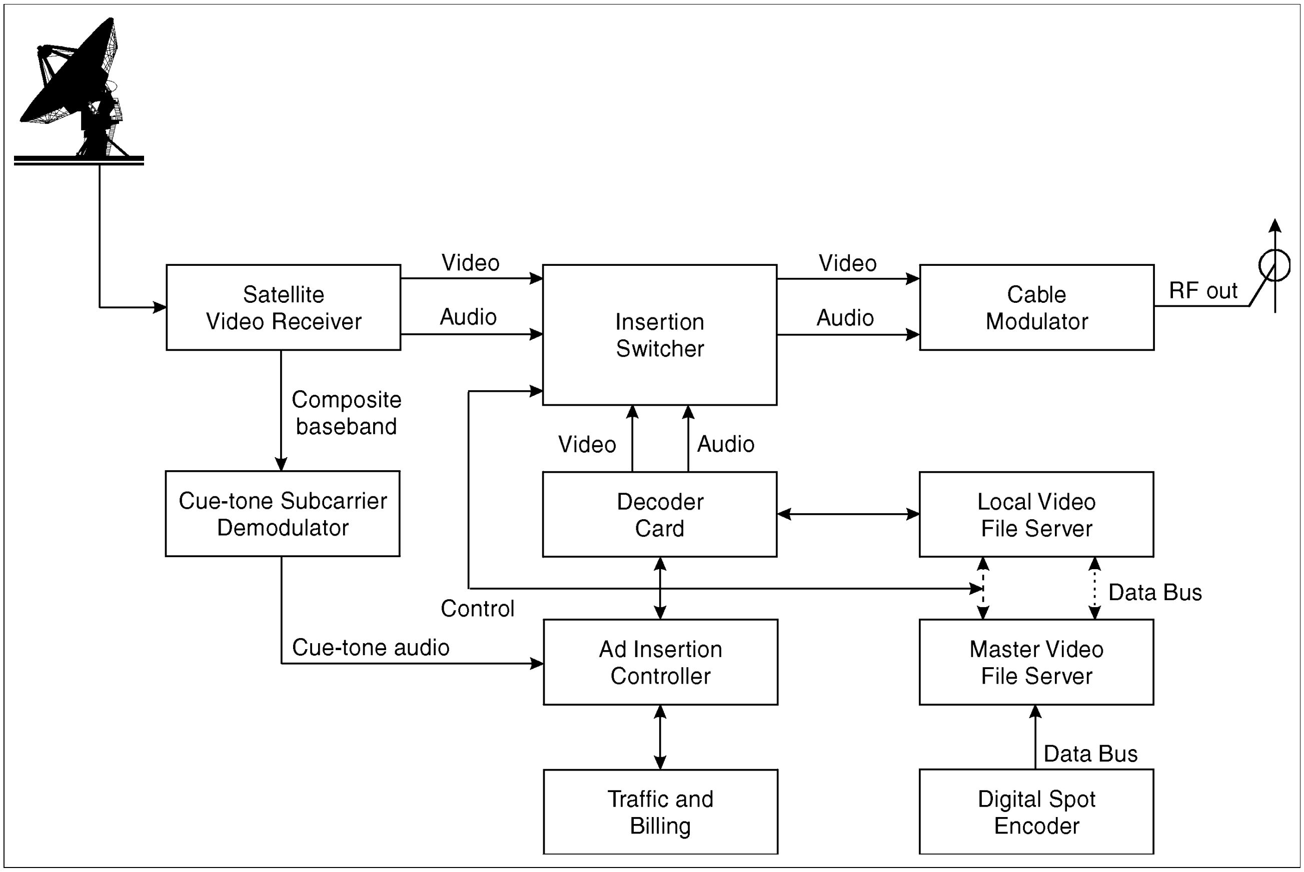

Figure 1. Hybrid Insertion System

modifications. Although operators are adding digital channels to their headends, a true digital ad insertion system is not commercially available. Currently, the added digital channels are used mostly for pay-per-view content. Very few commercials are inserted locally into those digital channels. Commercial insertion devices and systems will become more important as the transition from analog to digital program delivery progresses.

Advanced digital technology brings efficiency, flexibility, and other benefits over its analog counterpart. However, in some applications (*e.g.*, splicing MPEG-2 bitstreams) it can present more problems. Splicing is the fundamental technique used to insert commercials or short programs in channels, in editing audio/video content in post-production houses, and channel switching in headends and other broadcast stations.

ANALOG PROGRAM INSERTION

In the analog domain, splicing between two NTSC video sources or TV programs is relatively easy. A switch between two video sources at any frame boundary can be accomplished, since the frames are equal in size and time duration and are independent of one another. Analog insertion maintains synchronization among video sources prior to splicing. The resultant video, with the switched video source at the insertion or splice point, will appear seamless.

Figure 1 shows the block diagram of a typical commercial insertion system in cable headends and broadcast stations. In existing analog systems, networks distribute their content to cable headends and broadcast affiliates via satellite using a baseband bandwidth close to 10 MHz. As shown in Figure 2, the lower portion of the transmitted signal spectrum is used for NTSC video. The upper portion of the spectrum above 4.2 MHz contains a few FM subcarriers which are used to transmit mono, stereo, multi-language audio, and data. One of the subcarriers is used to transmit cue-tone signals. The subcarrier is modulated with a cue-tone signal using frequency modulation (FM) or frequency shift keying (FSK). Some networks send cue-tones using a pair of FSK tones (*e.g.*, 25 Hz and 35 Hz) within one of the audio channels. The integrated receiver-decoder (IRD) receives the RF-modulated signal, demodulates the cue-tone subcarrier signal, and provides an output signal. Either a dual-tone multiple frequency (DTMF) audio cue-tone sequence and/or a relay contact closure signal is provided to the insertion equipment. The networks send their schedules and spot availability (avail) in advance for local advertising insertion in a program. The precise location of the spot is signaled by a cue-tone during program broadcasts. The cue-tone consists of a sequence of numbers indicating the start and the end of an insertion opportunity. For example, the Discovery Channel uses a DTMF local avail cue-tone with 826* indicating the start, and 826# indicating the end of the spot.

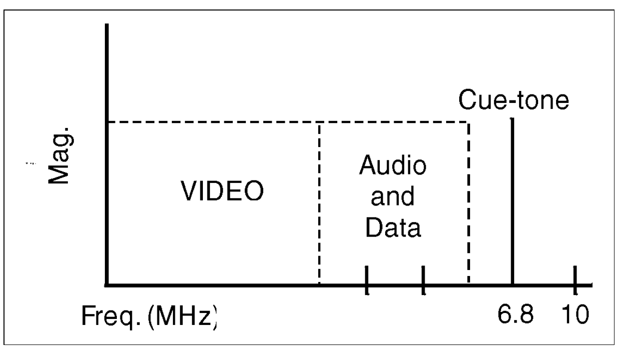

Figure 2. Analog Baseband Used by Networks

DIGITAL PROGRAM INSERTION

Compressed digital video frames are unequal in size (number of bits) due to the variable length coding compression techniques employed. This complicates locating the video frame boundaries. Not all frames in MPEG are independent of one another since substantial compression is achieved using temporal prediction between successive video frames. Also, compression algorithms may reorder the sequence of frames. The MPEG-2 standard has additional complexities, such as encoder-decoder clock synchronization, and decoder buffer management.

In an MPEG-compressed video elementary stream, the size of the compressed frames (access units), in terms of the number of bits, varies considerably. The buffer in the decoder smooths out the changes caused by this compression characteristic, as shown in Figure 3. For example, the size of an intra-frame (I-frame) may be 100 kbits, while that of a predictive-frame (P-frame) and a bidirectionally predictive-frame (B-frame) could be 35 kbits and 25 kbits, respectively. Hence, the compressed frames take unequal amounts of time to travel from the encoder to decoder at a fixed bitrate. However, the output of the decoder produces uncompressed video frames (presentation units) of equal size and duration. The buffer removes the effects of the variation in the arrival of the compressed frames on the non-varying periodic video decoding and display.

The minimum buffer size for a decoder must be specified for the buffer model of an MPEG-2 profile (*e.g.*, Main Profile at Main Level requires 1.8 Mbits). The decoder has no control of the decoder buffer fullness. When MPEG bitstreams are encoded, there is an inherent encoder buffer occupancy at every point in time. The same buffer fullness is reflected exactly in the decoder buffer. The constant transmission rate of the compressed video bitstream fills the buffer and the video decoder removes variable size compressed frames every display period, as shown in Figure 3.

Any discontinuity caused in the original bitstream, by instantaneously switching to a new bitstream, can cause the decoder buffer to overflow at some time in the future. The buffer fullness determines the delay or, equivalently, the time data spends in the buffer.

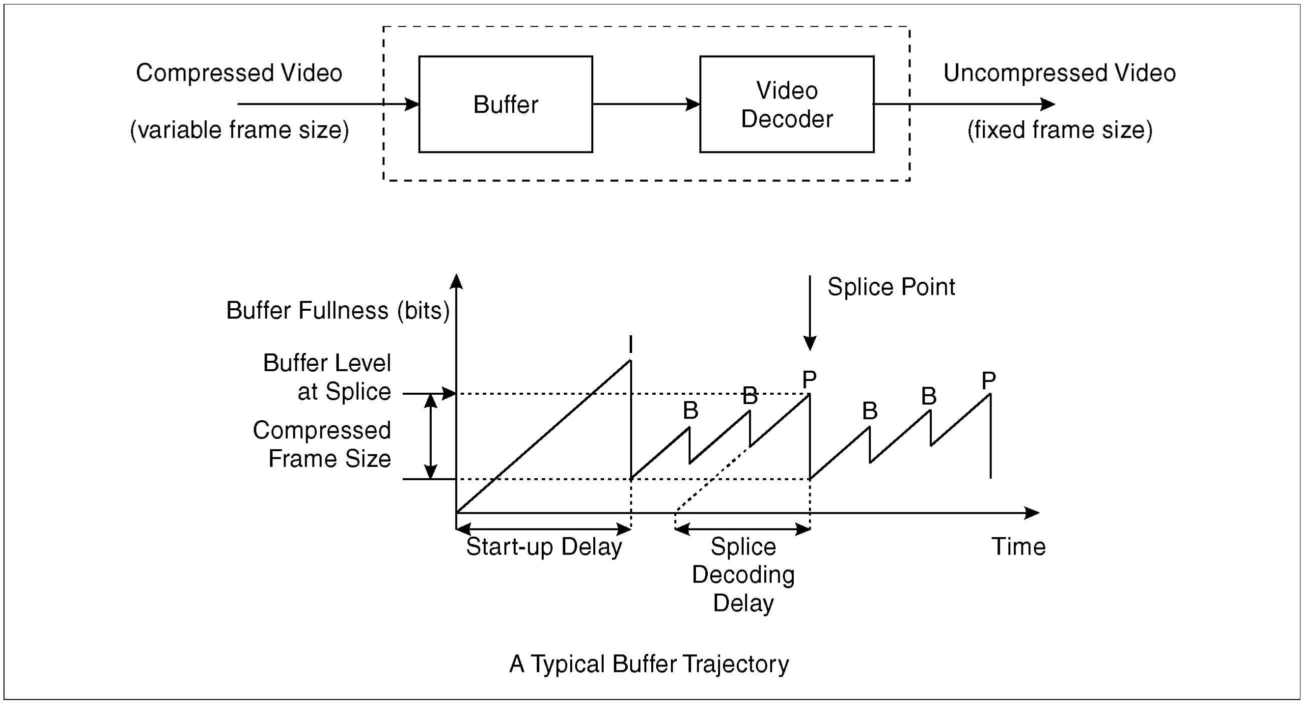

Figure 3. Simplified Representation of a Video Decoder

Seamless splicing requires that the encoder match the delay at the splicing point. Non-seamless splicing shifts the constraint from the encoder to the splicing device. The splicing device is responsible for matching the delay of the old and new streams as closely as possible without causing the buffer to overflow. Alternatively, controlled underflow of the decoder buffer, while displaying a repeated frame or several black frames before inserting the new stream, appears nearly seamless and prevents buffer overflow. Either splicing method may cause the audio buffer to overflow, resulting in an audio artifact. In most cases, this may be avoided by muting for a few audio frames.

As a result, splicing in the compressed domain is somewhat complex and needs additional care. Splicing points have to be identified in the bitstream at the time of encoding prior to the actual splicing operation. Without easily identifiable splice points, splicing would be more complex to implement. Some of the requirements include simplicity, low cost, and, above all, preserving current functionalities (such as cue-tone signaling) as much as possible. Identification of splice points is supported in the MPEG-2 standard, but operation of the splicing mechanism is not specified. Splicing can be implemented in a number of ways, deterring interoperable splicing devices. One of the important goals of standardization is to ensure that equipment is interoperable. This will allow cable operators to buy equipment from a competitive marketplace and not be tied to a single vendor. A program and commercial insertion system consists of three basic subsystems: insertion, billing, and trafficking. To standardize all these subsystems in the cable headend, the SCTE Digital Video Subcommittee (DVS) has created the Digital Program Insertion (DPI) *ad hoc* group.

The PT20.02 group of the Society of Motion Picture and Television Engineers (SMPTE) created an *ad hoc* group to devise a splicing standard. Similar efforts have also been undertaken by the Society of Cable Telecommunication Engineers (SCTE), the European Broadcast Union (EBU), and the Association of Radio Industries and Businesses of Japan (ARIB).

CableLabs has sponsored the Advertising in TV (Ad TV) group among its member cable companies to provide requirements and operational information to equipment manufacturers. CableLabs has contributed to the SCTE DVS group regarding some functionalities, features, and parameters desired for the next generation digital program insertion system. This contribution summarizes some of the comments received by CableLabs in response to the *Request for Information on Digital Program Insertion Systems* issued in April 1997. The comments address the following specific areas:

Bitrate

Many responding companies have voiced the need for mandatory support of constant bitrate insertion. Of the responses that also supported variable bitrate, most stated that the amount of bitrate variation required upper and lower bounds. Some suggested that an industry bitrate guideline be reflected in an insertion standard. Other responses suggested that the insertion system should have the capability to reduce the inserted program bitrate through recoding. In this case, some loss of video quality may result.

Audio

Mandatory support for Dolby AC-3 was suggested by many companies. There was also a request that the sampling rate be constant between old content and new. If audio and video splice points do not coincide, the audio splice point should be selected after the video splice point. Momentary muting of audio during the splice may be an attractive method to conceal audible artifacts occurring near the splice point. The audio of the new stream should be the first byte of an audio frame.

Addressability

Although many vendors stated their intent to support geographic and demographic addressability, they were concerned about substantial complexity and increased cost associated with adding this functionality. Low revenue-generating channels could be turned into more profit-generating channels by offering targeted advertising.

Emergency Messages

Mandatory support for emergency messages was confirmed by many companies. Consideration of FCC requirements about emergency messages should be taken into account where applicable.

Timing Reference

Some responses proposed the use of a global timing reference for scheduling the splicing operation. This constraint would require the cooperation of all program providers.

Interoperability

Most of the companies acknowledged the need for standardization to facilitate interoperability, one of the important goals of the cable industry. The responses generally included detailed physical interfaces that utilized formal and *de facto* standards. Consideration of interfaces to proprietary analog insertion systems was suggested. Such interfaces could allow replacement of “core” splicing technology and associated control within existing analog insertion systems. There were requests to integrate existing standards for various interfaces to take advantage of economies of scale.

To facilitate interoperability at the subsystem level, CableLabs proposed a Logical Topology for the next generation DPI system in their Request For Information (RFI) of April 1997. This architecture, shown in Figure 4, has been modified and submitted to both the SCTE DVS and SMPTE groups.

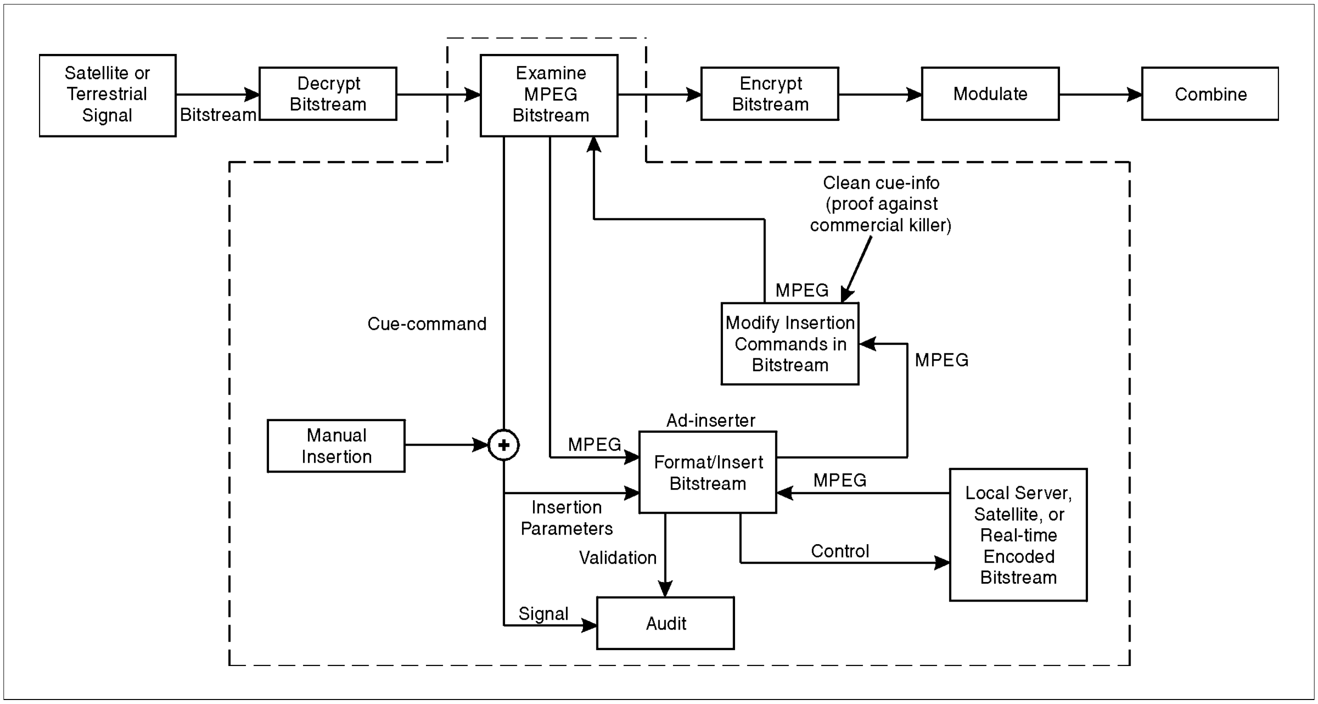

Figure 4. Typical Architecture of a Digital Program Insertion System

INSERTION SOLUTIONS

SMPTE also collected requirements from other application sectors, primarily in the television production environment. Two of these sectors, studios and post-production houses, have been influential in designing the new SMPTE standard. Production facilities prefer seamless splicing, whereas cable headend and broadcast affiliates prefer a less complex (non-seamless or near seamless) inexpensive design.

The SMPTE splicing method switches between two video streams in the compressed domain. To minimize complexity while reducing stream overhead and cost, SMPTE made major assumptions which constrain the compressed bitstreams for splicing as follows:

Splicing will be performed between unencrypted streams (implying splicing will be performed in the clear).

1. The streams must have equal bitrates.
2. The streams must have identical raster formats (*i.e.*, an equal number of pixels per line).
3. All legacy decoders (decoders already deployed) must be transparent to splicing for commercial insertion and other purposes.
4. Splicing will leave no signature in the post-spliced stream to alert potential commercial killers.
5. A cue-tone signaling mechanism (analogous to existing analog methods) has been included to indicate a specific splicing point, start and duration of a single avail, schedule of avails over a period of time, and other information within the MPEG-2 Transport Stream for one or more programs.

The SMPTE splicing method switches between two video streams in the compressed domain. As a result, some of the disadvantages of the SMPTE splicing standard are:

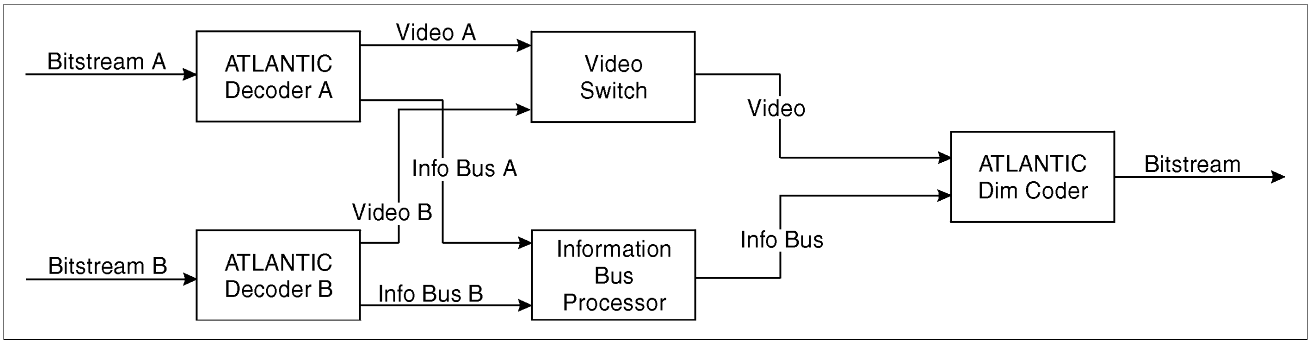

Figure 5. Simplified Diagram of an ATLANTIC Video Switch

1. Identification of splice points has to be inserted at the encoder, including additional buffer management, with a possible concurrent loss in compression efficiency.

2. Splicing cannot be achieved at every video frame boundary—a feature inherent in the analog method—because of the dependency of the compressed frames on each other.

3. Encoding of a splice point requires an increased number of bits. Increasing the number of splice points in a bitstream increases overhead and effectively increases the bitrate.

Similarly, the European Broadcast Union (EBU) has been working on projects involving the interface and manipulation of MPEG-2 compressed bitstreams. One of their efforts addresses compressed video bitstream switching and editing under the ATLANTIC project. As previously mentioned, switching MPEG-2 bitstreams is complex, and splicing of compressed bitstreams does not offer the flexibility required for all switching applications. The ATLANTIC project proposes a solution based on the decoding of the MPEG bitstreams with all mixing, switching, or other production operations carried out on uncompressed video streams prior to re-encoding into a new compressed bitstream.

Naive decoding of the compressed bitstream to uncompressed video, and then re-encoding (recoding) back, degrades video quality due to cascaded lossy compression. Using a lower compression factor on the material to be switched and re-encoded can reduce these artifacts. A disadvantage of this approach is that the initial compression requires a high bitrate and, therefore, would not be appropriate for efficient broadcast delivery.

The ATLANTIC approach uses a partial decoding technique. To minimize degradation in the output bitstream, the coding decisions (temporal predictions encoded as motion vectors, quantization, and information in the picture header), used in compressing the sequence of frames, are directly used or passed through, when possible, in the recoding process. As shown in Figure 5, the uncompressed decoded video streams are routed to video switches, while saved coding decisions are sent to the information bus processor. A few frames around splice points may change temporal prediction modes, and motion vectors may need to be recalculated in these instances only for a few frames. However, a vast majority of the frames may remain unchanged and the recoder can use the motion vectors available from the info_bus processor. Therefore, recoding is much less complex than full encoding, and no noticeable degradations in quality are observed even after several decodings and recodings at the same bitrate.

The ATLANTIC project proposes a method of mixing the information in the info_bus (the so-called MOLE™) with the decoded uncompressed digital D1 signal to enable unconstrained post-production effects and editing. Therefore, once the bitstream is decoded, the initial encoding decisions are embedded in the uncompressed stream. This information will be reinserted during re-encoding. The same stream could be decoded and re-encoded a number of times using a simpler and less complex encoder without noticeable degradation in picture quality. The EBU has submitted a proposal to SMPTE for standardizing MOLE™ technology so that interfaces for standard production equipment, enabling transparent production and editing of compressed video bitstreams, may be standardized and manufactured by various vendors. The real advantage in this method is its flexibility and lack of constraints. No splice points have to be created prior to splicing, and splicing can be performed at any frame boundary. In fact, any type of production effect can be achieved since editing is accomplished in the uncompressed domain.

The authors believe that this technology is attractive and has many potential applications. However, recoding is substantially more complex than operating solely on compressed bitstreams. Therefore, implementing this technology with the components available today may be more expensive. An implementation using customized VLSI processors (currently not available), such as media processors, may be more economical and, therefore, more widely available.

CONCLUSION

As mentioned earlier, splicing and switching standardization efforts are in progress. SMPTE and EBU have standardized the basic splicing of MPEG-2 bitstreams in two different approaches with different constraints, flexibility, and complexity trade-offs. Efforts are also underway to standardize the use of basic splicing in digital program/commercial insertion and its other integral aspects, such as trafficking and billing. The SCTE DVS DPI *ad hoc* group and the AdTV (Advertising in TV) group are working on these two areas. Still, a few issues remain to be addressed by the cable industry, such as changing the original compressed bitrate of the ad spot, ad insertion into a statistically multiplexed variable bitrate stream, and use of constrained bitstreams per the SMPTE method or unconstrained streams per the ATLANTIC method.

Today's average channel bitrate is 3 to 4 Mbits/sec (sports up to 6 Mbits/sec). For quality reasons, the bitrate of the compressed commercials may be higher than the average channel bitrate of the program targeted for ad insertion. The question remains—how to insert a higher bitrate ad in a lower bitrate channel? For efficient use of the spectrum, it is better to keep the channel multiplex unaffected. In this case, a trade-off may be required to code (or recode if already compressed at higher bitrate) the ad at the same channel bitrate with a corresponding reduction in picture quality.

In statistical multiplexing, the channel bitrates specified are the average rate and the maximum and minimum limits of the bitrate variation. If an ad is encoded at a constant bitrate, the variable rate of the statistically multiplexed channel targeted for insertion has to be maintained at a constant matching rate during insertion. This may affect the multiplex efficiency and, consequently, the video quality of the other channels.

For basic splicing, the SMPTE method is simpler than the ATLANTIC method. But the SMPTE method requires several bitstream constraints, and splicing cannot be performed at all frame boundaries. The ATLANTIC method is more flexible, does not have any bitstream restrictions on frame boundaries, and bitrate reduction can be accommodated. However, it is more complex than the SMPTE method. Advanced programmable processors may make this implementation more cost-effective by sharing economies of scale with other applications.

The cue-command mechanism provided in the SMPTE splicing standard preserves and extends the cue-tone features available today. Designing a device to detect cue-commands in a transport multiplex should be straightforward.

In summary, several operational issues need to be resolved that impose constraints on the delivery of compressed programming targeted for the insertion of local advertising. Relaxation of some of these constraints requires a more complex solution. Standardization is needed for interoperable digital program insertion equipment from multiple vendors. If the cable industry can achieve consensus on these issues, such standardization will facilitate the timely introduction of such new digital commercial insertion systems.

ACKNOWLEDGMENTS

The authors would like to thank Don Morgan and Carol Derr of TCI, Joe Davis of Digital Video, Reem Safadi of GI, and Steve Riedl of Sky Connect for many useful discussions and information, as well as their ongoing participation in the various digital program insertion technology development and standardization activities.

REFERENCES

1. *Cable Advertising*, white paper, March 1997, Cable Television Laboratories, Louisville, CO.

2. *Cable Ad Sales, Business Goals, Objectives, Strategy and Technical Vision*, November 1992, Cable Television Laboratories, Louisville, CO.

3. *Digital Program Insertion, Request for Information*, April 28, 1997, Cable Television Laboratories, Louisville, CO.

4. *Splice Points for MPEG-2 Transport Streams*, SMPTE Standard, December 1997.

5. *Seamless Concatenation—A 21st Century Dream*, M.J. Knee, Snell Wilcox Limited, U.K., N.D. Wells, BBC Research and Development, U.K., September 1997.