RECOMMENDATION ITU-R BR.1356*

User requirements for application of compression in mainstream standard definition television production** and archival***

(Questions ITU-R 238/11 and ITU-R 239/11)

(1998)

The ITU Radiocommunication Assembly,

considering

a) that new disk-based storage supports are expected to penetrate all areas of television production, namely non-linear editing, on-air playout and short- to medium-term archives;

b) that this technology offers significant gains in terms of vastly improved operating flexibility, production flow and station automation and is therefore highly attractive for the up-grading of existing studios and the design of completely new studio installations;

c) that the economic and time efficient use of hierarchical storage technology together with new interconnecting networks which allow interactive and multi-user operation mandates, however, a significant reduction of bit rate of the video signal subjected to the above processes;

d) that a number of mutually incompatible bit-rate reduction schemes to achieve economic data storage as well as different file formats and networking protocols for signal interchange have been already introduced to the market which jeopardize interoperation between individual equipment and remote studios of different manufacture;

e) that studies of compression schemes for television data storage and archival are deemed to be particularly important and urgent and will be beneficial to broadcasters, including those in developing countries, as it is demonstrated by contributions received from ITU-D, the WBU and some administrations,

recommends

1 that compression algorithms and transport schemes should be based on open standards. This implies availability to all interested parties on a fair and equitable basis of the intellectual

---

* Radiocommunication Study Group 6 made editorial amendments to this Recommendation in 2001 in accordance with Resolution ITU-R 44.

** Mainstream production refers to production/post-production that aims at achieving quasi-transparent performance at a compressed video bit rate around 50 Mbits/s.

*** The scope of this Recommendation is limited to 525-line and 625-line interlaced systems ans video compression schemes using bit rates of 50 Mbit/s or less (excluding audio) and a coding resolution of one TV-frame or longer.

A general tutorial description of compression algorithms suitable for use within television production as well as some application examples are given in Appendix A to this Recommendation.

property necessary to implement those standards. Availability in the marketplace of chip sets and/or algorithms for software encoding and decoding may give users confidence in the adoption of particular compression methods;

2 that the number of compression methods and parameters should be minimized for each uniquely defined television production application in order to maximize compatibility and interoperability;

3 that compliance testing methods should be available for those building equipment to standards for algorithms and transport schemes and for users purchasing and installing equipment to those standards. Standards bodies should adopt standards for compliance testing methods to support both manufacturer and user needs;

4 that a single compression scheme used with different compression parameters throughout the chain should be decodable by a single decoder;

5 that the development of a common ("agile") decoder is desirable to support the use of more than one compression family;

6 that integration of video compression into more complex systems must be via standardized interfaces. Translating through Recommendation ITU-R BT.601, i.e., decoding and re-encoding, should be the default method of concatenating video signals compressed using different techniques and/or parameters;

7 that the compression scheme chosen should not preclude the use of infrastructures based on the serial digital interface (SDI) as embodied in Recommendation ITU-R BT.656;

8 that issues related to interoperability must be further explored and standards developed to allow predictable levels of performance to be achieved in the implementation of specific applications;

9 that bit streams carrying compressed signals should be designed so that they can be formatted and packaged for transport over as many types of communications circuits and networks as possible;

10 that appropriate channel coding methods and error protection be employed where necessary;

11 that compression systems should be designed so that, in normal operation, signal timing relationships (e.g. audio/video lip sync) and synchronization presented at encoder inputs are reproduced at decoder outputs;

12 that signal delays through compression processing (encoding/decoding) must be limited to durations that are practical for specific television production applications;

13 that provision should be made for selected analogue vertical interval information to be carried through the compression system, although not necessarily compressed with the video. Provision should be made to carry selected parts of metadata through a transparent path synchronously with the video and audio data;

14 that the compression scheme chosen for devices that mimic VTRs should allow for the reproduction of pictures in shuttle mode for identifying content and of pictures in jog and slow motion modes for selecting edit points;

15 that network interfaces and storage devices should provide for both variable bit rate (VBR) and constant bit rate (CBR) options;

16 that storage devices should allow recording and playing back of television programme streams and files as data rather than decoding to baseband for recording and re-encoding upon playback;

17 that the compression strategy chosen for standard television should be extensible to high definition applications to allow for commonality in the transitional phase.

NOTE 1 – Specialized terms frequently used in the context of innovative television production are listed in Recommendation ITU-R BR.1357.

APPENDIX A

Report on the use of compression in television production

Introduction

Digital compression for video represents the core enabling technology for innovative television programme production of the future since it permits the storage of programme material in servers that make access to that material virtually instant and allows simultaneous use by multiple users. These features all have the effect of improving workflow efficiency and reducing the cost of producing, post-producing and repurposing of television programmes.

The report identifies a number of parameters which define the basic performance of different compression schemes and enlarges on their impact on picture quality and post-production headroom within a variety of applications normally encountered in distributed television production.

PART A

Compression issues

1 Digital compression for video

1.1 Image quality

Selection of compression system parameters has a significant impact on overall image quality. These compression parameter choices must be optimized to preserve image quality while at the same time fitting the image data into the available bandwidth or storage space. Different combinations of compression parameters may be best for different specific applications.

Compression system parameters which should be considered include: the underlying coding methods, the coding sampling structure, pre-processing, data rates and the group of pictures (GOP) structure used. In choosing compression system parameters, interaction between parameter choices must also be considered. Finally, special operational issues such as editing the bit stream or splicing new content into an incoming bit stream should be considered.

1.1.1 Coding method

The coding method is the most fundamental of compression choices. There are four main compression methods used in the television production and distribution chain: MPEG-2 Main Profile at Main Level (MP@ML), MPEG-2 4:2:2 Profile at Main Level (4:2:2@ML), Motion JPEG, and DV. All of these coding methods are based on the discrete cosine transform (DCT). All of these coding methods use normalization and quantization of the transform coefficients, followed by variable length coding.

MPEG includes motion estimation and compensation in its tool kit of techniques. This allows improved coding efficiency, with some cost penalty in memory and processing latency. Motion JPEG and DV are both frame-bound, thereby minimizing coding cost, but these frame-bound coding methods do not take advantage of the coding efficiency of inter-frame motion estimation and compensation. MPEG and DV both allow motion adaptive processing in conjunction with intra-frame processing.

1.1.2 Sampling structure

MPEG, Motion JPEG, and DV can all be used with the 4:2:2 pixel matrix of Recommendation ITU-R BT.601. MPEG and Motion JPEG can both be used with other pixel matrices, multiple frame rates, and either interlace or progressive scan. Note that the 4:2:2 matrix is sub-sampled from the original full-bandwidth (4:4:4) signal. The pixel matrix can be further sub-sampled to reduce the signal data, with 4:2:2 sampling normally used for interchange between systems.

4:2:2 systems, such as the MPEG-2 4:2:2 Profile, 4:2:2 Motion JPEG, and the DV 4:2:2 system (which operates at 50 Mbits/s), all use half the number of colour difference samples per line compared to the number used in the luminance channel. 4:2:2 provides half the horizontal bandwidth in the colour difference channels compared to the luminance bandwidth and full vertical bandwidth.

4:1:1 systems such as DV 525 use one quarter the number of colour difference samples per line compared to the number used in the luminance channel. 4:1:1 reduces the colour difference horizontal bandwidth to one quarter that of the luminance channel while maintaining full vertical bandwidth. The filters used to achieve the 4:1:1 sub-sampled horizontal bandwidths, like other horizontal filters, generally have a flat frequency response within their pass-bands thereby enabling translation to and from 4:2:2 with no further degradation beyond that of 4:1:1 sub-sampling.

4:2:0 systems such as DV 625[1] and MPEG-2 Main Profile, use half the number of colour difference samples horizontally and half the number of colour difference samples vertically compared to the number used in the luminance channel. 4:2:0 therefore retains the same colour difference horizontal bandwidth as 4:2:2 (i.e., half that of the luminance channel) but reduces the colour difference

[1] Note that the proposed SMPTE D-7 format, although based on DV coding, will use 4:1:1 sampling for both the 525- and 625-line systems.

vertical bandwidth to half that of the luminance channel. 4:2:0 coding, however, generally does not provide flat frequency response within its vertical pass-band, thereby precluding transparent translation to the other coding forms. Consequently, systems that use 4:2:0 sampling with intermediate processing generally will not retain the full 4:2:0 bandwidth of the prior coding.

Care must be exercised in selecting compression sampling structures where different compression coding techniques will be concatenated. In general, intermixing different sub-sampled structures impacts picture quality, so cascading of these structures should be minimized. For example, while 4:1:1 or 4:2:0 signals will have their original quality maintained through subsequent 4:2:2 processing (analogous to "bumping up" of tape formats), cascading 4:1:1 and 4:2:0 will generally yield less than 4:1:0 performance.

1.1.3 Compression pre-processing

Video compression systems have inherent limitations in their ability to compress images into finite bandwidth or storage space. The compression systems rely on removal of redundancy in the images, so when the images are very complex (having very little redundancy), the ability to fit into the available data space may be exceeded, leading to compression artefacts in the picture. In these cases, it may be preferable to reduce the complexity of the image through other methods before compression processing. These methods are called pre-processing, and include filtering and noise reduction.

When noise is present in the input signal, the compression system must expend some bits encoding the noise, leaving fewer bits for encoding the desired image. When either motion detection or motion estimation and compensation is used, noise can reduce the accuracy of the motion processing, which in turn reduces coding efficiency. Even in compression systems which do not use motion estimation and compensation, noise adds substantial energy in high frequency components of the DCT which might otherwise be zero. This not only wastes bits on extraneous DCT components, but degrades run length coding efficiency as well.

Compression system specifications generally define only the compression functions within equipment, but do not specify the pre-processing before the compression function. An exception is the shuffling which is an inherent part of the DV system, and is not to be confused with the shuffling used for error management in digital recorders.

Since most pre-processing, such as filtering or noise reduction, is not always required, the pre-processing parameters may be selected depending on the nature of the images and the capabilities of the compression system. These choices can be pre-set or can be adaptive.

1.1.4 Data rate

The MPEG-2 4:2:2 Profile at Main Level defines data rates up to 50 Mbits/s. Motion JPEG 4:2:2 equipment typically operates at data rates up to 50 Mbits/s. DV 4:1:1 and DV 4:2:0 operate at 25 Mbits/s. DV 4:2:2 operating at 50 Mbits/s is undergoing standardization. MPEG-2 Main Profile at Main Level is defined at data rates up to 15 Mbits/s.

Selection of the data rate for MPEG-2 4:2:2@ML is interrelated with the group of pictures (GOP) structure used. Lower bit rates will typically be used with longer, more efficient, GOP structures, while higher bit rates will be used with simpler, shorter GOP structures. Intra-coded images

(MPEG-2 4:2:2 Profile I-pictures only, M-JPEG, and DV), at data rates of 50 Mbits/s, can yield comparable image quality. MPEG-2 4:2:2@ML with longer GOP structures and lower data rates can provide comparable quality to shorter GOP structures at higher data rates – albeit at the expense of latency. (See group of pictures below.)

1.1.5 Group of pictures

There are three fundamental ways in which to code or compress an image. The most basic is coding a field or frame with reference only to elements contained within that field or frame. This is called intra-coding (I-only coding for short). The second way in which to code an image uses motion compensated prediction of a picture (called a P-picture) from a preceding I-coded picture. Coding of the prediction error information allows the decoder to reconstruct the proper output image. The third method also uses motion compensated prediction, but allows the prediction reference (called an anchor frame) to precede and/or follow the image being coded (bidirectional or B-picture coding). The selection of the reference for each picture or portion of a picture is made to minimize the number of bits required to code the image.

Sequences of images using combinations of the three coding types, as defined by MPEG, are called groups of pictures (GOPs). Both Motion JPEG and DV use only intra-frame coding and therefore are not described in terms of GOPs.

MPEG-2 allows many choices of GOP structures, some more commonly used than others. In general, a GOP is described in terms of its total length and the repetition sequence of the picture coding types (e.g. 15 frames of IBBP). The optimal choice of GOP structure is dependent on the specific application, the data rate used, and latency considerations.

Since I-only pictures are least efficient and B-pictures are most efficient, longer GOPs with more B- and P-pictures will provide higher image quality for a given data rate. This effect is pronounced at lower data rates and diminished at higher data rates. At 20 Mbits/s, the use of long GOPs (e.g. IBBP) may prove useful, while at 50 Mbits/s shorter GOPs can provide the required quality.

Besides affecting image quality, the choice of GOP structure also impacts latency. Since a B-picture cannot be coded until the subsequent anchor picture is available, delay is introduced in the coding process. Note, however, that this delay is dependent on the distance between anchor frames, not the total length of the GOP structure. This means that a blend of the coding efficiency of long GOP structures together with the lower latency of short GOP structures can be obtained by judicious use of P-picture anchors.

1.1.6 Constant quality vs. constant data rate

Compression systems are sometimes referred to as variable bit rate (VBR) or constant bit rate (CBR). MPEG-2 and Motion JPEG can operate in either variable bit rate or constant bit-rate modes; DV operates only with constant bit rate. In practice, even those systems commonly believed to be constant data rate have data rate variation, but over shorter periods of time. Another way to characterize the compression systems is to compare constant quality with constant data rate systems.

Constant quality systems attempt to maintain a uniform picture quality by adjusting coded data rate, typically within the constraint of a maximum data rate. Since simpler images are easier to code, they are coded at lower data rates. This results in more efficient compression of simpler images and can be a significant advantage in storage systems and in non-real-time transfer of images. Constant quality operation is useful for disk recording and some tape recording systems such as tape streamers.

Constant data rate systems attempt to maintain a constant average data rate at the output of the compression encoder. This will result in higher quality with simpler images and lower quality with more complex images. In addition to maintaining a constant average data rate, some constant data rate systems also maintain the data rate constant over a GOP. Constant data rate compression is useful for videotape recording and for fixed data rate transmission paths, such as common carrier services.

Constant data rate processing will, of course, be characterized by a target data rate. Variable data rate processing can be constrained to have a maximum data rate. By ensuring that this maximum data rate is less than the target rate of the constant data rate device, constant quality coding can operate into a constant data rate environment.

1.1.7 Editing

Consideration of compression parameters relating to editing falls into two general applications categories: complex editing and simple cuts-only editing (seamless splicing). In the case of complex editing involving effects or sophisticated image processing and analysis, many of the processes will require decoding back to the Recommendation ITU-R BT.601 domain. In these cases, the coding efficiency advantage of complex GOP structures may merit consideration. In the case of cuts-only editing, however, it may be desirable to perform the edits entirely in the compressed domain using bit stream splicing. Bit-stream splicing can be done between two bit streams which both use the same compression method. Data rates and other parameters of the compression scheme may need to be bounded in order to facilitate splicing. Some existing compressed streams can be seamlessly spliced (to provide cuts-only edits) in the compressed domain with signals of different data rates.

Techniques for operating directly in the compressed domain are still being developed. Issues relating to editing in the compressed domain are being addressed. It has even been suggested that carrying out more complex operations in the compressed domain may be possible. It should be noted, however, that much of the image degradation encountered in decompressing and re-compressing for special effects will similarly be encountered if those effects operations are performed directly in the compressed domain, since the relationships of the DCT coefficients will still be altered by the effects.

If all the compression coding methods used in an editing environment are well defined in open standards, systems could include multi-format decoding. Multi-format decoding would allow receiving devices to process compressed streams based on a limited number of separate compression standards, thereby mitigating the existence of more than one compression standard.

1.1.8 Concatenated compression

To the extent possible, television systems using video compression should maintain the video in compressed form rather than employing islands of compression which must be interconnected in uncompressed form. Since several compression and decompression steps are likely, the ability to withstand concatenated compression and decompression is a key consideration in the choice of a compression system. The results of concatenated compression systems will be influenced by whether the systems are identical or involve differing compression techniques and parameters.

There are a number of factors which influence the quality of concatenated compression systems. All the systems considered here rely on the DCT technique. Anything which changes the input to the respective DCT operations between concatenated compression systems can result in the transformed data being quantized differently, which in turn means that additional image information is lost. Further, any changes which result in different buffer management will result in different quantization.

In the case of MPEG coding, any change in the alignment of GOP structure between cascaded compression steps will result in different quantization, since the P- and B-picture transforms operate on motion compensated image predictions, while the I-picture transforms operate on the full image.

For MPEG, M-JPEG, and DV, any change in the spatial alignment of the image between cascaded compression steps will result in different quantization, since the input to any particular DCT block will have changed. Any effects or other processing between cascaded compression steps will similarly change the quantization.

Concatenated compression processes interconnected through Recommendation ITU-R BT.601 will have minimal image degradation through successive generations if the compression coding method and compression parameters, including spatial alignment and temporal alignment, are identical in each compression stage.

It is not always possible to avoid mixing compression methods and/or parameters. In some applications the total image degradation due to cascaded compression and decompression will be minimized by attempting to maintain the highest quality compression level throughout and only utilising lower quality compression levels where occasionally necessary, such as in acquisition or in using common carrier services. For other applications, however, which must make greater use of lower quality compression levels, the best overall image quality may be maintained by returning to the higher compression quality level only where dictated by image processing requirements.

Beyond the quality issues just discussed, there are operational advantages realized by staying in the compressed domain. Faster than real-time transfers, as well as slower than real-time transfers, can be facilitated in the compressed domain. Further, some users would welcome image processing in the compressed domain as a potential means to achieving faster-than-real-time image processing.

1.2 Quality levels

While different compression performance levels will be used in different application categories, users will attempt to minimize the total number of performance levels within their operation. Performance differences will be accompanied by differences in cost of equipment and of operation that are appropriate to the application category. For example, a typical broadcast operation might have three levels of compression quality.

The highest compression quality level, generally requiring the highest data rate, would be used in applications which require the highest picture quality and in applications which involve extensive post-production manipulation. A key attribute of this quality level is the ability to support multiple generation processing with little image degradation. The highest compression quality level might therefore be used in some higher-quality production applications, but production applications which require the very highest quality will continue to use uncompressed storage and processing. The highest compression quality would also be used for critical imagery and to archive programme content which is likely to be re-used in conjunction with subsequent further production processing.

A middle compression quality level would be used in applications which require good picture quality and in applications which involve some limited post-production manipulation. This quality level would support a limited number of processing generations, and might be used for news acquisition, news editing, network programme distribution, and local programme production. The quality level would also be used to archive programme content which may be re-used but is not likely to involve significant additional production processing.

A lower compression quality level would be used in applications which are more sensitive to cost than to quality. This quality level would not normally support subsequent processing but might be used for programme presentation or mass storage for rapid-access browsing. The lower compression quality would not generally be used to archive programme content which might be reused.

These examples of highest, middle, and lower compression quality levels do not necessarily correspond to any particular absolute performance categories, but rather should be taken as relative quality levels to be interpreted according to the specific requirements of a particular user's criteria. Further details on particular applications and their use of compression can be found in Part B – Compression applications.

1.3 Operational considerations

Systems of all compression performance levels must be fully functional in their intended applications. Equipment employing compression should function and operate in the same or better manner as similar analogue and non-compressed digital equipment. The use of compression in any system should not impede the operation of that system.

If it is possible to select and alter compression characteristics as part of the regular operation of a compressed system, such selection and alteration should be made easy by deliberate design of the manufacturer. Variable compression characteristic systems should possess user interfaces that are easy to learn and intuitive to operate. In addition, selections and alterations made to a compressed system must not promote confusion or compromise the function and performance of systems connected to it.

More than a single compression method or parameter set might be employed in a television production facility. Where this is the case, these should be made interoperable. Compression characteristics used in the post-production process must concatenate and inter-operate with MPEG-2 MP@ML for emission.

It is well recognized that integration of compressed video systems into complex systems must be via standardized interfaces. Even with standardized interfaces, however, signal input/output delays due to compression processing (encoding/decoding) occur. System designers are advised that this compression latency, as well as stream synchronization and synchronization of audio, video, and metadata must be considered. Efficient video coding comes at the cost of codec delay, so balance must be achieved between minimum codec delay and the required picture quality. This may be particularly important for live interview feeds, especially where the available bandwidth is low and the real time requirement is high. Compressed systems must be designed to prevent the loss of synchronization or disruption of time relationships.

Compressed signal bit streams should be designed so that they can be formatted and packaged to permit transport over as many communications circuits and networks as possible. Note that compressed bit streams are very sensitive to errors and therefore appropriate channel coding methods and error protection must be employed where necessary.

Provision should be made for selected analogue vertical interval information to be carried through the compression system, although not necessarily compressed with the video. Additionally, selected parts of the ancillary data space of digital signals may carry data (e.g. metadata), and provision should be made to carry selected parts of this data through a transparent path synchronously with the video and audio data.

1.4 Storage

Where a compressed video bit stream is stored and accessed on a storage medium, there may be storage and compression attributes required of the storage medium depending on the intended application.

1.4.1 Interfaces

Recommendation ITU-R BT.601 is the default method of interfacing. However, as network interfaces become available with the required guaranteed bandwidth access and functionality, they will allow methods of digital copying between storage devices. Because storage devices can both accept and deliver data representing video at non-real time, the network should also allow the transfer of files at both faster and slower than true time for greater flexibility. The network interface should allow the options of variable bit rate (VBR or constant quality) and constant bit rate (CBR) at different transfer bit rates, and optionally transfer specialized bit streams optimized for stunt modes. This will allow a downstream device to copy a file directly from a primary device for stunt mode replay on the secondary device.

1.4.2 Data rate requirements

Where possible, users would prefer to record incoming data directly as files on a data storage device rather than decoding and re-encoding for storage. Since there will be different compressed video bit rates depending on the application, any network connection to the device must be capable of a wide variety of input and output data rates.

Both a tape streaming device and a disc based video server will need to be able to store variable bit rate compressed video streams. This will require an interface which can accommodate the requirements of a VBR data stream.

Furthermore, compressed video streams may be stored on a tape streamer or disc server with each stream recorded at a different average bit rate.

1.4.3 Resource management

A tape streamer needs to be able to accept and present compressed video files over a range of values. An integrated system will need to know how to control the streaming device for an I/O channel which may have a programmable data rate rather than a constant data rate.

The storage devices should specify the range of data rates which can be recorded and played back. A disc based video server additionally has the capability of accepting multiple I/O channels. Further signalling may be necessary to ensure that both channel bandwidth and the number of channels can be adequately signalled to the system.

1.4.4 Audio, video and metadata synchronization

Many storage devices may record video data, audio data and metadata on different parts of the media or on separate media for various reasons. Synchronization information should be included to facilitate proper timing of the reconstructed data at normal playback speed.

1.4.5 VTR emulation

Where a storage device using compressed video is intended to be or to mimic a VTR, it may implement VTR stunt modes. Such stunt modes may include viewing in shuttle mode for the purpose of identifying the content, pictures in jog and slow motion for the purpose of identifying editing points, and broadcast-quality slow-motion. However, the removal of redundancy from the video signal by compression will naturally reduce the possibilities for high quality stunt mode reproduction. Compression methods and parameters must allow stunt mode capability where required in the user's application. If the recording device is required to reconfigure the data onto the recording media to provide better stunt mode functionality, such conversion should be transparent and impose no conversion loss.

1.5 Interoperability

Interoperability can be a confusing term because it has different meanings in different fields of work. Compression systems further confuse the meaning of interoperability because of the issues of programme transfers, concatenation, cascading, encoding and decoding quality, and compliance testing. Programme exchange requires interoperability at three levels: the physical level, the protocols used, and the compression characteristics. This chapter considers only compression while other chapters address physical links and protocols.

Considering Programme transfers, the Task Force has identified that there are several types of interoperability. The first example identified is inter-operation through Recommendation

ITU-R BT.601 by decoding the compressed signals to a raster and re-encoding them. This is the current default method and is well understood. Additional methods of inter-operation are expected to be identified in the future. Further work for this Task Force is to categorize methods of inter-operation, explore their characteristics and relate them to various applications, and develop possible constraints on device and system characteristics necessary to ensure predictable levels of performance sought by users for specific applications.

1.6 Compliance testing

Interoperability between compressed video products is essential to successful implementation of systems using compression. Although inter-operation is possible via Recommendation ITU-R BT.601, it is desirable to have inter-operation at the compressed level to minimize concatenation losses. Compressed inter-operation can involve encoders and decoders using the same compression method and parameters, the same compression method with different parameters, or even different compression methods. Compliance testing is a fundamental step toward ensuring proper interoperability.

Compliance testing can be employed by manufacturers and users of compression systems in a variety of ways. Encoders can be tested to verify that they produce valid bit streams. Decoders can be tested to verify that a range of compliant bit streams can be properly decoded. Applications can be tested to verify that the characteristics of a given bit stream meet the application requirements, for example whether the amount of data used to code a picture is within specified limits. In practice, defining and generating compliance tests is more involved than applying those tests, so tests employed by manufacturers might be identical to tests employed by users.

In the case of MPEG-2, compliance testing focuses on the bit stream attributes without physical compliance testing, since MPEG-2 does not assume a particular physical layer. A number of standardized tests are described in ISO/IEC 13818-4. The concepts for tests specified in the MPEG-2 documents may be extended to other compression methods, including Motion JPEG and DV. These compliance tests include transport stream tests, programme stream tests, timing accuracy tests, video bit stream tests, and audio bit stream tests. The MPEG-2 video bit stream tests include a number of tests specific to the 4:2:2 Profile at Main Level. Other tests will need to be developed.

PART B

Compression applications

1 The application model

Figure 1 illustrates a generic view of the programme chain identifying all the primary application areas. Each of these functions is described in detail below.

2 Acquisition

Compression offers considerable benefits for field acquisition, e.g. in news, sports, electronic field production (EFP), consumer, and allied uses.

The requirements for equipment for field application are:

– high portability;

– low weight;

– low power consumption (long running time); and

– ruggedness.

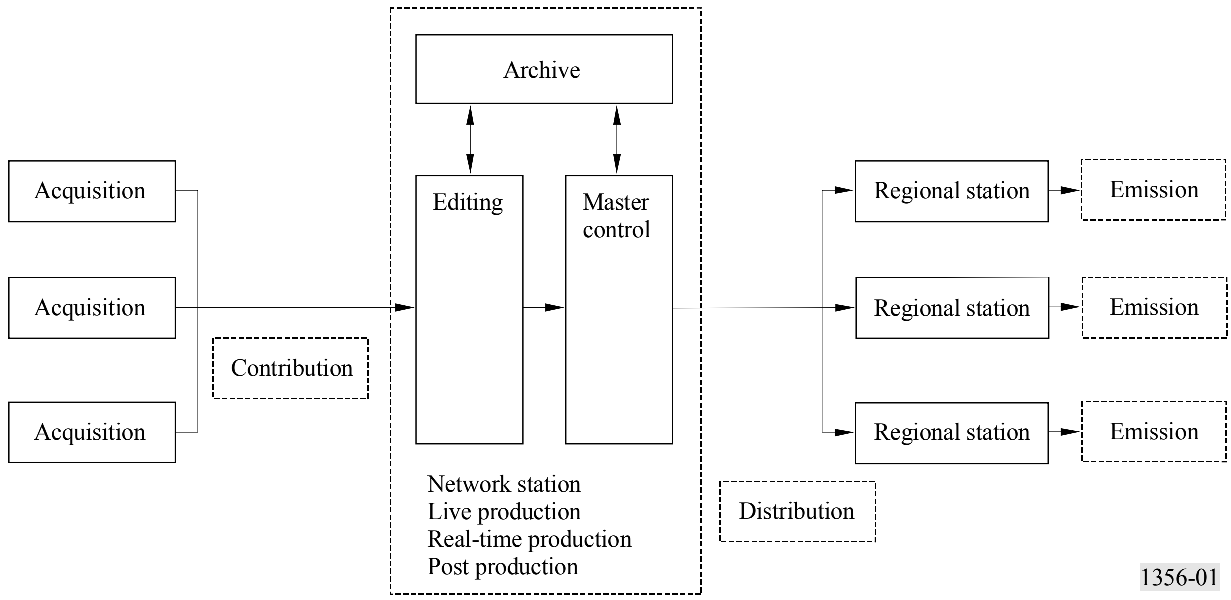
The application of compression plays a significant role in meeting these requirements because of the reduced bit rate and reduced need for storage capacity that it yields.

The compression scheme used for acquisition should allow field editing of the signal.

For applications such as EFP and Sports applications a 4:2:2 sampling structure is preferable. Usage of a sampling structure of less than 4:2:2 (such as 4:1:1 and 4:2:0) should be acceptable for News Gathering applications provided extensive post-processing is not required.

FIGURE 1

Generic television programme chain



3 Contribution

Content acquired in the field and edited on the spot using a portable field editor or fed live must be delivered to the studio for further processing and final transmission.

Communication channels available for this purpose may have a different bandwidth than the recorded compressed signal and possibly a different quality of service. It would be desirable for the compressed signal from an acquisition source to fit perfectly into the constraints of the communication medium. However, due to the wide variety of choices in the marketplace (e.g. DS3, E3, and OC-1 through OC-192, with data rates ranging from tens of megabits to gigabits per second) this cannot be practically achieved without resort to the data rate of the lowest bit bandwidth channel.

Where a compression system has a bit rate greater than that of the communications channel, the options are slower-than-real-time transmission (which may not be practical for some applications) or an additional layer of compression, which will lead to concatenation errors and will result in some degradation of picture quality.

In the case where the compression system used has a data rate lower than the communications channel it will be preferable to use a single compression system rather than use concatenated compression systems. The unused bandwidth may be used for other purposes (e.g. extra error protection or ancillary service information).

Satellite contribution service between remote trucks and studios is not always available due to overcrowding of the spectrum for this type of communication. A choice of low bit-rate compression may be helpful in this communication shortfall. However, using this as the only criterion for selection of the bit rate of the compressed signal may have other negative consequences.

Significant input to television news is received in the form of fixed satellite communication. This type of a signal is recorded at the station on a 24-hour basis and selection of the proper compression format is dependent on other criteria which may be different from those of field acquisition equipment.

4 Compression in live/real-time production

Use of compression in field production will ease demand on the communication channel between remote acquisition equipment locations and the studio where final production is completed. The application of compression in live and real-time production has been covered in the sections on acquisition and contribution.

5 Post-production/exploitation

5.1 Editing in real time

Editing in real time usually involves the simplest of production effects – the cut, implemented by stream splicing in compressed systems. The splice should be done in the compressed domain without transcoding.

It is highly desirable to be able to perform cuts editing on compressed streams with different bit rates. This area of compression processing is currently in its infancy. However, it is becoming clear that this will be possible between streams of the same compression family.

5.2 Post-production editing

Post-production editing requires the highest attainable quality. During such production processing, layering is very often used. Any processing effects beside cuts require decoding. Layering, as a process, may involve many subsequent re-encodings; therefore loss of picture quality can be a real possibility.

To avoid this loss of picture quality phenomenon, the producer should use methods similar to those used in the past for layering in analogue environments. One of these methods involved the use of higher quality processing and storage formats to ensure that the original signal is degraded as little as possible. Extensive studies have been performed on concatenation of compressed systems and these studies have shown that concatenated compression will have a negative impact on overall picture quality. The analogue practice of editing and using intermediate storage having the highest quality format available is also applicable in the compressed environment. This means that the processing should be done in the uncompressed domain or in the highest compressed bit rate available to the user.

5.3 Off-line post-production

It is assumed that for off-line post-production, the user has less demanding picture quality requirements because this process is being used only for creation of an edit decision list (EDL). The EDL is later used for actual editing of the programme master. In this manner, a post-production facility can utilize high quality, high cost post-production equipment efficiently. An important feature of the off line editing station is its ability to facilitate proper editing decisions which can be accurately followed in the on-line environment.

5.4 Presentation/master control

This type of processing controls the output stream of the station or network. In this environment decisions are always made live and usually under dynamic conditions. The operator must make decisions based on specific situations and therefore response time and repeatability are of the highest importance. At this processing point, the operator has little control over picture quality, only over delivery of picture content.

Reliability of the equipment is also of the utmost importance, and any possibility of failure could have catastrophic consequences for the facility. It is clear that the use of compression will not improve this situation. However, if a facility operates with compressed signals as its primary methodology, it is logical that master control should follow the processing pattern of the facility and switch compressed signals.

Switching in master control will take place at the packetized elementary stream (PES) level. This process may be followed by encoding into transport streams. Nevertheless, the output picture quality will have been predetermined by previous encoding processes, and the final picture quality will be decided by the bit rate and structure of the released compressed stream.

6 Distribution

6.1 Satellite links

Satellite links are frequently available as analogue channels within which a suitable digital modulation process can be utilized. This allows considerable flexibility in the choice of video compression system bit rates. Where input bit streams have been previously encoded, satellite encoders should use the same coding process as that of the feeder content, or the pre-encoded content should be directly passed to the satellite modulator to ensure no concatenation losses. Where this is not possible (because the available bit rate is less than that of the pre-encoded bit stream) concatenation losses will be incurred.

6.2 Public carriers (telcos)

The issue of public carriers has already been covered in the section on "Contribution" where the same points apply.

Other features of both satellite links and Telco links include the possibility of faster-than-real-time transfers for both cost and operational benefits.

6.3 Emission

Transmission to the home is subject to low bit-rate limitations and likely to be MPEG-2 MP@ML, probably limited to less than 8 Mbit/s but certainly less than 15 Mbit/s.

Insertion of local content may be required in some circumstances through the use of compressed stream "splicing".

6.4 Packaged media distribution

Distribution of content to the home is, like emission, subject to low bit-rate limitations. A major difference, however, is likely to be the use of the VBR technique and a single channel output.

Insertion of local content will not be required, but provision for interactive branching within the stored content will, in some cases, be necessary.

7 Archiving

7.1 On-line archiving

On-line archives will generally directly record compressed bit streams to avoid the concatenation effects of another compression system. The archive may also be associated with highly compressed browse mode images and metadata to aid recovery of archive content. The quality level of the browse mode images is only required to support picture recognition, with no expectation that these pictures will be used for any other purpose.

Metadata and browse mode images will normally be located on the same storage device for rapid access to the content. Full editing capability should be possible with on-line archive content.

7.2 Near-line archiving

A near-line archive is a mid-way archive containing copies of the content and metadata. Typically, browse mode images and metadata will still be stored in the on-line archive for rapid access whereas content will be stored off-line on a remote server. Full editing capability of near-line archive content is still possible.

7.3 Deep/long-term archiving

In a deep archive, content and all associated metadata will be stored in the archive, although only the metadata and browse mode images may be readily accessed for archive browsing. Depending on the deep archive requirements, the content may be subjected to a further compression process such as MPEG-2 MP@ML using a low bit rate.