Transparent concatenation of MPEG compression

N. Wells

BBC Research & Development

The techniques described here allow the MPEG compression standard to be used in a consistent and efficient manner throughout the broadcast chain.

By using a so-called "MOLE" which is buried within the decoded programme material, it is possible to concatenate (i.e. cascade) many MPEG encoders and decoders throughout the broadcast chain – without any loss of audio or video quality.

The described techniques have been developed in the ATLANTIC Project [1] which is a European collaborative project within the ACTS framework.

1. Introduction

The MPEG compression standard[1] will be used for the distribution of many new digital TV services. Also, MPEG compression is already being used for contributions into the studio, because of bandwidth/bit-rate restrictions on some incoming connections. In addition, there will be pressure to use high levels of compression in future TV archives in order to give on-line access to thousands of hours of programme material. MPEG-2 compression would be a sensible choice for such archives as this standard gives a video compression performance which is difficult to improve upon, given the likely requirements for quality and bit-rate, and for the broad range of picture material to be archived [2].

However, once the signal has been compressed into MPEG form, it becomes difficult to perform operations on the signal of the sort normally encountered along the production and distribution chain. For example, it is not possible to edit or switch simply between two MPEG bitstreams without causing serious problems for a downstream decoder. Ideally, we would like to be able to handle and operate on the compressed signal in just the same way that we handle the PAL/NTSC signal today. Inevitably, this requires that the signal is decoded before being passed through "traditional" mixing or editing equipment and then re-coded at the output of the process. Then, however, more than one generation of compression has been applied to the signal. Along the complete production and distribution chain, it is likely that the signal will undergo several generations of decoding and re-coding. With multiple generations of

---

1. In this article, "MPEG" is used to mean MPEG-2 MP@ML video compression and MPEG-1 Layer II audio compression.

compression, the picture and sound quality can degrade very rapidly as the number of generations increases.

This degradation of quality can be avoided by intelligent re-coding or "cloning" of the MPEG signals after decoding. The techniques described here open up the possibility of MPEG being used for post-production and all stages of distribution at bit-rates little different from those used for the final broadcasting stage.

2. The production chain

A simplified model of a typical programme production and broadcasting chain for a future MPEG digital TV service is shown in *Fig. 1*.

Within the studio of *Fig. 1*, a single programme is assembled from local sources and possibly from archive or satellite contribution material that has already been coded in MPEG form. Programme assembly will involve switching, mixing and editing of the various contributions. This can only be realistically achieved by working with uncompressed/decoded signals in the standard studio format, since it is important to be able to mix between material that exists in a number of different source

**Figure 1
Model of an MPEG
broadcasting chain.**

formats (e.g. tape, servers, live inputs etc.). At the output of the studio, the final programme will be assembled and compressed to MPEG form with the inclusion of several elements in addition to the main audio and video components. These elements might include subtitles (closed captions), multiple sound channels and references to Web pages etc. All associated signals and data are synchronized with the main audio and video components via the MPEG syntax.
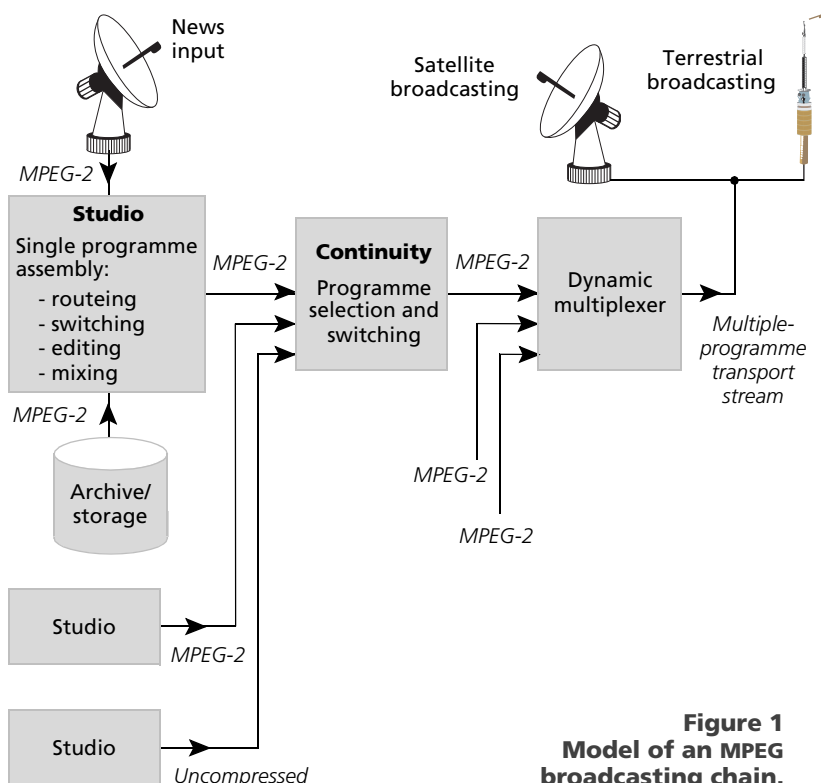
The Playout or "Continuity" Centre shown in *Fig. 1* is responsible for ordering and scheduling the output of a given network channel, and for adding links and inserts between individual programmes. The most convenient format for the input bitstream to Continuity will probably be MPEG because of all the additional components associated with a given programme. However, programmes may be delivered to Continuity in many different compressed and uncompressed formats. Again the only feasible way to switch and mix between different programme material is in the decoded domain. After Continuity, the continuous channel output will be compressed into a continuous MPEG bitstream for multiplexing together with other bitstreams into a multiple-programme stream.

The final channel output may be distributed over more than one network (e.g. satellite and cable) and there may well be a requirement to change the bit-rate of the signal in accordance with the requirements of each separate network. In order to change the bit-rate of an MPEG signal in an optimum way, some degree of decoding and then re-coding is required.

In addition to the elements shown in *Fig. 1*, there could well be a requirement for the insertion of local programmes into a nationally-distributed bitstream. In this case, one programme item is removed from the national multiplex and is replaced by a locally-derived programme item. This effectively repeats the "Continuity" function and involves a further decoding and re-coding of the associated channel.

Consequently, along the programme production and distribution chain, the signal might easily encounter up to five cascaded encodings and decodings and this could lead to severely degraded picture and sound quality. What is required is a solution that enables a signal to be decoded and then re-encoded without the build-up of compression impairments. The solution developed within the ATLANTIC project is based around the "MOLE" as described in the next section. MOLE-based techniques were first proposed in [3].

3. Introducing the MOLE [2]

3.1. Video

3.1.1. Transparent cascading

It is possible to decode a video signal from MPEG and recompress it back to an almost identical MPEG bitstream (a clone of the first bitstream), provided that the second encoder can be forced to take exactly the same coding decisions as were taken by the first encoder. This is not necessarily an obvious result because the input to the second encoder contains coding noise introduced into the source signal by the first coding and decoding process. A short explanation which illustrates how the transparency of decoding followed by re-coding can be achieved is given in the adjacent text box.

The relevant decisions/parameters used by the first encoder which must be re-used in the second encoder include the following:

⇨   the motion vectors for each macroblock;

⇨   the prediction mode for each macroblock (frame/field, intra/non-intra, forward/backward/bi-directional etc.);

⇨   the DCT type for each macroblock (frame/field);

⇨   the quantization step size for each macroblock;

⇨   quantization weighting matrices.

These parameters are necessarily carried within the syntax of an MPEG bitstream because they are required by a decoder to decode the bitstream. What is required is a method of conveying these parameters along with the *decoded* video.

The method being proposed by the ATLANTIC project is to bury the information invisibly in the video signal itself. The buried information signal is called a "MOLE". A straightforward

2.   This term has been protected as a Trade Mark by one of the ATLANTIC partners.

method for carrying the MOLE is to use the least significant bit (10th bit) of the chrominance component in the standard digital interface for component video signals (ITU-R Recommendation 601). Three factors which support this format for the MOLE are:

⇨ the data is invisible even on the most critical test material;

⇨ MPEG is basically an "8-bit format" and therefore the two least significant bits of the standard 10-bit interface are not active for a signal that has been decoded from MPEG-2;

⇨ subsequent (8-bit) encoders will not code this chrominance bit.

It should be noted that, in order to be able to generate the MOLE, no additional information has to be added to the bitstream apart from that required to decode the bitstream.

3.1.2. MOLE-based architecture

A basic video switch/mixer architecture using the MOLE is shown in *Fig. 2*. It comprises a standard component digital mixer with inputs coming either from MPEG decoders or from an uncompressed source such as a camera or from some other form of digital decoder such as a JPEG decoder. The MPEG decoders add the MOLE information to their decoded output. When a decoded MPEG input is selected by the mixer then the decoded signal plus the MOLE is carried transparently through the mixer to the following MOLE-assisted encoder. 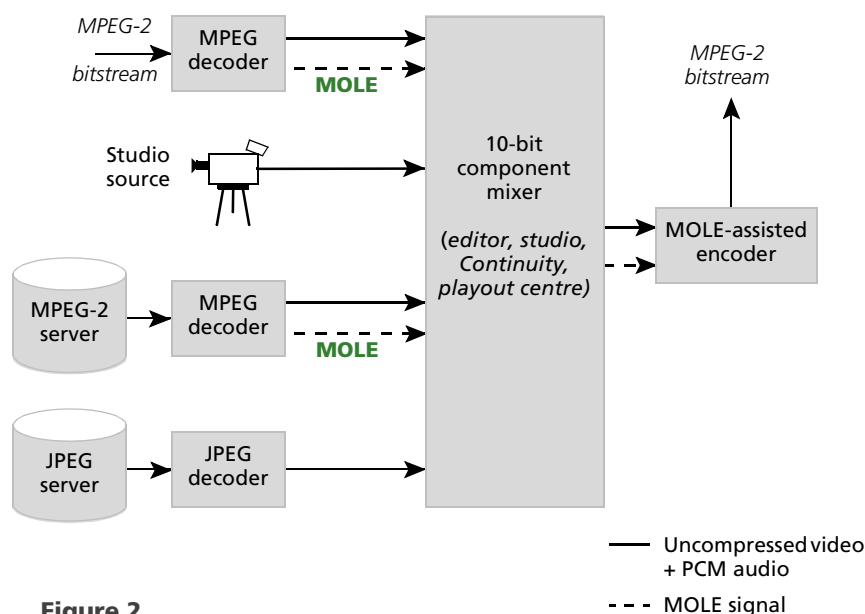This encoder recognizes that a MOLE is present and locks its own internal decision processes to the parameters carried in the MOLE. Then the output MPEG bitstream will be the same as the selected input MPEG bitstream.

Figure 2
MOLE-based switching/mixing.

During a switch or cross-fade to another decoded MPEG input on the digital mixer, there will be some frames where the MOLE signal is not valid or has become corrupted. The MOLE signal contains information which enables checking of the validity or corruption of the information carried; if the MOLE is not valid, then the encoder uses its own internally-derived parameters in place of those carried in the MOLE. When the switch or cross-fade has been completed and the second decoded MPEG signal has passed transparently through the mixer, then the MOLE signal will again become valid and the encoder can lock onto the new information. Within a few frames the coder will be producing an MPEG bitstream which is the same as that being fed to the second decoder.

Consequently, such an architecture provides for a seamless transition from one MPEG bitstream to another. This is achieved without imposing any constraints on the type or relative timing of the Group of Pictures (GoP) structures of the input MPEG bitstreams, nor any constraints on the frames at which the transition occurs. Away from the transition there is no

Abbreviations

ATM	Asynchronous transfer mode		**JPEG**	(ISO) Joint Photographic Experts Group
CBR	Constant bit-rate		**MAP**	*Maximum a-posteriori*
CRC	Cyclic redundancy check		**MCP**	Motion-compensated prediction
DCT	Discrete cosine transform		**MPEG**	(ISO) Moving Picture Experts Group
DSM-CC	(ISO) Digital storage media command control		**PCM**	Pulse code modulation
EDL	Edit decision list		**PES**	Packetized elementary stream
ETSI	European Telecommunication Standards Institute		**SMPTE**	(US) Society of Motion Picture and Television Engineers
GoP	Group of pictures		**TCP**	Transmission control protocol
HDTV	High-definition television		**VBR**	Variable bit-rate
IDCT	Inverse discrete cosine transform		**VLC**	Variable-length coder
ISO	International Organization for Standardization		**VLD**	Variable-length decoder
IT	Information technology		**VTR**	Video tape recorder

loss of quality resulting from the cascaded decoding and re-coding of the MPEG bitstreams. However, during the transition period, the signals are effectively decoded, combined and re-coded with new coding parameters (such as picture type and quantizer step size etc.). Simulations and initial real-time tests of such a switching process have consistently shown that any generational loss of picture quality is not visible during the short period of the transition [4].

Because the switching is done in the decoded domain, this architecture enables MPEG compression to be used without loss in conjunction with conventional systems which use no compression or only mild compression (such as the Digibeta, JPEG, DV or SX formats). When the MPEG source is selected, the signal will be re-coded without loss because of the presence of the MOLE. When a non-MPEG source is selected, the MOLE will cease to be valid and will then disappear. At this point the coder will start to use its own internally-generated decisions to move seamlessly towards coding the new source signal as a stand-alone coder.

A MOLE-based architecture can be used equally well with MPEG-2 video bitstreams which have been coded in a variable bit-rate (VBR) mode, and with bitstreams which have been coded in a constant bit-rate (CBR) mode.

3.1.3. Video MOLE format

A format for the MOLE has been proposed and is currently under discussion for standardization within the EBU/ETSI Joint Technical Committee and the SMPTE [5].

In the proposed format, the MOLE data is both picture- and macroblock-locked; this means that the data which relates to a given 16-pixel by 16-line macroblock is co-sited with these 256 pixels

on the 10th bit of the chrominance samples in the macroblock. Of the available 256 bits per macroblock, the majority of these are taken up with data that changes at macroblock rate, e.g. the motion vector data. Information that only changes at the picture rate is distributed across the picture in reserved slots within the macroblock data format. This picture-rate information is repeated five times across the picture in case some parts of the picture are changed during the mixing operations.

Other information carried in the MOLE data includes a rolling macroblock count and a cyclic redundancy check (CRC) across all the data in the macroblock. The macroblock count is not picture-locked and can be used to detect a wipe or switch between two different decoded sequences. The CRC is used to detect whether the MOLE data has been corrupted as a result of any picture processing applied to that macroblock. In order to reduce any possibility of the MOLE data being visible, the data is scrambled using a method known as "signalling in parity". The parity of one chrominance sample (including the MOLE bit) and the following luminance sample is made "odd" to carry a data bit equal to "1" and made "even" to carry a "0" data bit.

3.1.4. Examples of MOLE in use

A particular example of the use of MOLE data is in the insertion of captions or logos into a decoded MPEG sequence. Those macroblocks within a picture which have been changed in any way by the inserted caption or logo can be detected by using the CRC data. The coder can then re-code the affected macroblocks using locally-derived optimum decisions. Those parts of the picture which are not affected by the insertion can be re-coded transparently using the valid MOLE data.

The MOLE should also be applicable in cases where the original MPEG sequence was coded with fewer pixels per (active) line than the number defined for the digital studio standard. For example, some early MPEG implementations for standard-definition TV chose to code only 704 out of the standard 720 pixels/line. Alternatively, the MPEG signal may have been coded at a lower horizontal sampling frequency such as 528 samples/line. In such cases, after decoding to the full studio standard of 720 pixels/line, it should be possible, if required, to re-code back to the same MPEG bitstream with the same number of samples/line and with the macroblocks in the same positions relative to the picture material. Therefore, it is necessary for the MOLE data to include some form of synchronization code which can be used to locate the positions of the original macroblocks in the decoded data. Note that when a lower horizontal sampling frequency has been used, the area corresponding to a coded macroblock in the decoded (and up-sampled) picture has a length greater than 16 pixels. Also, when a lower horizontal sampling frequency has been used, it is necessary for the process of up-sampling followed by down-sampling of the video to be transparent. This can be done by using a carefull combination of up- and down-filters for sample-rate conversion to and from the full sample rate.

3.1.5. Alternative methods for carrying MOLE data

In some cases, it may not be appropriate to carry the MOLE data on the least significant bit of the decoded chrominance component; for example, it may be required to store the decoded MPEG sequence on a video tape recorder which uses a small degree of compression. This compression would be sufficient to corrupt the MOLE data without perhaps adding any visible degradation to the picture material. In this case, the MOLE information can be carried as

N. Wells

an ancillary signal. An efficient way to code the MOLE information is then to keep the data in pseudo-MPEG-2 form but to remove all the video coefficient information (which takes up most of the bit-rate in a typical MPEG-2 bitstream).

3.1.6.   Chrominance subsampling

The version or type of MPEG-2 coding which will be used primarily for distribution is referred to as "Main Profile". In order to obtain the best overall picture quality at a given bit-rate, this profile uses half the vertical chrominance sampling frequency of the studio standard (e.g. 4:2:0 as opposed to 4:2:2 resolution). Therefore, each coder is required to vertically pre-filter the chrominance component before reducing the sampling rate prior to coding, and each decoder is required to vertically filter the chrominance output as it increases the vertical chrominance sampling rate back to the full rate.

In the cascaded decoding and re-coding process shown in *Fig. 2*, it is possible that the cascaded application of up- and down-conversion filters adds further resolution loss to the chrominance component. However, it is easily possible to make the system transparent to the up- and down-conversion processes by ensuring that the combined response of the decoder and encoder filters is "Nyquist". The presence (or not) of a MOLE can be used to determine whether or not the video signal has undergone any previous filtering and can be used to adapt the coder pre-filter accordingly.

3.2. Audio

The same MOLE ideas can be applied to audio in order to avoid the impairments introduced by succesive decoding and re-coding of compressed audio signals. Such cascading is inevitable in the TV broadcast chain shown in *Fig. 1* but it is also likely to occur in similar audio-only production and distribution chains for digital audio broadcasting.

For transparent decoding and re-coding, the second coding process is required to take the same coding decisions as the initial coder. For audio, the main decisions which need to be kept constant are (i) the positions of the audio block boundaries and (ii) the quantization step sizes for each of the frequency sub-bands within each block. For MPEG Layer-II coding, the block boundaries occur at regular intervals; for example, at 24 ms intervals for 48 kHz sampling. A quantization step size is transmitted in the compressed bitstream as a combination of two parameters, namely a *scale factor* and a *bit-allocation* for the sub-band.

As with the video, the audio MOLE information can be added to the least significant bit of the decoded PCM audio signal; for example the 20th bit in typical digital audio installations. It is proposed [6] to scramble the MOLE data via "signalling in parity" whereby the MOLE data is used to control the parity of each (20-bit) audio PCM sample. A 20th-bit MOLE is completely inaudible and even a 16th-bit MOLE (for 16-bit audio PCM) is only just perceptible on the most critical material under carefully-controlled listening conditions.

Information carried by the audio MOLE for MPEG Layer-II coding includes the following:

⇨   block synchronization word;

⇨   number of bits of MOLE data per frame;

⇨   an indication of the original sampling frequency;

⇨   mode information (mono, joint stereo etc.);

⇨   copy and copyright flags;

⇨   timing offset information;

⇨   error-checking bytes.

The timing offset information listed above is included primarily for use in TV switching and editing. This field carries information about any "lip-sync" error which may have been introduced during a switch because of the requirement to have both video frame continuity and audio frame continuity in the switched bitstream. Because the audio and video frames have different periods, it will be necessary to advance or delay the audio (by up to 12 ms for Layer-II) in relation to the video after a switching point. The timing offset information can be used to prevent such delays from accumulating along the broadcast chain.

The audio MOLE allows MPEG audio bitstreams to be switched and edited using conventional digital audio studio equipment which may be part of a TV or radio production chain. However, if the audio signal is processed in any way (remote from the switching point) then the MOLE will be corrupted. This means that the gain or frequency equalization of the audio signal should not be altered if transparent transcoding is required. Such a constraint is traditionally more acceptable in TV production than in radio production. If it is required to change the audio signal in some way then transparent cascading is not possible; but quality can be conserved in many circumstances by taking account, in the re-coding, of the MOLE information which would then have to be sent via an auxiliary data path.

4.   Changing the bit-rate (transcoding)

There will be a requirement along the TV production and distribution chain to change the bit-rate of the MPEG-2 signal. In particular this will apply to the video component of the signal which occupies the major part of the bit-rate of any single programme. The rate may be changed for example across the playout/continuity mixer shown in *Fig. 2* when the input MPEG bitstream is sourced at a higher bit-rate than that required for distribution.

Within an MPEG encoder, the average bit-rate is determined by the coarseness of the quantization applied to the DCT coefficients. When there is no change in rate on re-coding, then the quantizer in the re-coder does not introduce any further change in the value of the DCT coefficients (see the text box on page 13). However, when the bit-rate is changed, then a second stage of quantization must be applied to the DCT coefficients, thus introducing further noise into the signal. This noise can be minimized by exploiting the knowledge obtained through the MOLE about the quantizer in the previous generation of coding. An optimum quantizer, specifically for transcoding, has been designed within the ATLANTIC project and is referred to as a "MAP" (maximum a-posteriori) quantizer [7][8]. The MAP quantizer specifies how ranges of input levels are mapped onto standard output levels defined in the MPEG standard. This mapping is based on a parametric model of the impairments introduced by the previous generation of quantizer.

Also, by using information carried in the MOLE about the bit-rate statistics of the input bitstream, it is possible to define a good single-pass-rate controller for use within the second-generation encoder [9].

For transcoding, experiments have been done to compare the performance of various quantizers in the second-generation encoder. The results show that the MAP quantizer performs sig-

nificantly better than a quantizer that has been optimized for stand-alone, single-generation encoding [9]. Also, experiments have shown that, for an optimized two-stage coding (e.g. 5 Mbit/s to 3 Mbit/s), the subjective picture quality at the final bit-rate is no worse than that obtained in going from the source picture to the lower final bit-rate in a single generation, using a coder with a quantizer that is optimized for single-generation encoding.

This is an important result because it means that we are free to change the video bit-rate at critical points in the programme production and distribution chain without suffering any subjective quality penalty in the final decoded output. As a consequence, this allows the use of MPEG compression in archive storage and programme production at bit-rates which are slightly higher than those which might be currently required for distribution. This means that the picture quality/bit-rate of the archived material can be chosen to suit future as well as current requirements.

5. Editing and post-production

5.1. The MOLE and post-production

Using a MOLE-based architecture as shown in *Fig. 2*, it is possible to switch/mix between two MPEG bitstreams with no cascading loss, except for a small imperceptible loss close to the transition. The switching point can be specified to frame accuracy at any point within the GoP structure of the input MPEG bitstreams. Consequently, we have a system which can be used as the basis for editing MPEG bitstreams or for editing between MPEG bitstreams and formats that use other forms of compression (or no compression at all).

For the type of programme material that does not involve sophisticated picture manipulation during post-production, the acquisition and post-production could be done using MPEG at the bit-rate which will be used for final distribution. Alternatively, the bit-rate could be maintained at a slightly higher value and transcoded for final distribution. The advantages of using low bit-rate MPEG are:

⇨ low capacity servers;

⇨ low bandwidth servers;

⇨ low bandwidth networking.

For standard-definition TV, a typical bit-rate for an MPEG signal in post-production might be 8 Mbit/s or 1 Mbyte/sec. At such bit-rates it is possible to use conventional IT networks and servers for carrying the programme material. By contrast, other compression schemes being proposed for studio production use bit-rates up to 50 Mbit/s. In such cases, specialized networking solutions dedicated to these high bit-rates are required together with large and specialized servers.

The bit-rates proposed for these other compression schemes are high for two main reasons: (i) because they use little or no motion-compensated processing in order to give frame-accurate editing capability and (ii) to keep the quality high in order to avoid perceptible degradation with multi-generation cascading. However, the problems of frame-accurate editing and multi-generation cascading can be solved by a consistent use of MPEG and the MOLE throughout the production and distribution chain. This solution will be particularly relevant for economic post-production of HDTV because of the significantly higher bit-rates of HDTV signals.

5.2.   Small studio reference architecture

5.2.1.   Functional overview

For post-production, the ATLANTIC project has chosen to develop prototype equipment and applications according to the studio reference architecture shown in *Fig. 3*.

In this architecture, MPEG signals which arrive at the studio are passed through a *format converter* which separates the audio and video components and then packages these in a standard form (as MPEG "PES" packets with one "access unit" or "frame" per PES packet). These standard bitstreams are stored as files on the *main server* together with index files which relate time-code for a given frame to the corresponding byte location within the file of compressed data. The audio and video are separated because there is a requirement in many modern studios for "bi-media" working where radio production and TV production share the same studio and source material. In such studios it would make for inefficient use of network and server bandwidths if both the audio and video information had to be accessed just to get at the audio component.

Figure 3
Small studio architecture based around MPEG and ATM networking.

One disadvantage of the MPEG format in post-production is that it is not a particularly convenient format for browsing through data. This is because the coding algorithm uses inter-frame prediction which results in functions such as reverse play, fast-forward and fast-reverse that are rather limited in performance. Therefore, in the architecture of *Fig. 3*, as the MPEG files are placed on the main server, the signals are transcoded into a second format which is more suitable for browsing and determining the edit points; for example, this could be a browse-quality JPEG format as used in many conventional non-linear editors. In ATLANTIC, the browse format was chosen to be a low-resolution MPEG I-frame only, at a bit-rate of about 4 Mbit/s. The browse data is also accompanied by an index file which relates the timecode of each frame to its byte location within the browse file. The browse data may be stored on a separate *browse server*.

Edit decisions are then taken "off-line" using non-linear editors working with the browse data. The resulting edit decision lists (EDLs) are transferred to an "edit conformer" which is basically a MOLE-based switcher/mixer as shown in *Fig. 2*, but under automatic control. The EDL controls the fetching of data from the appropriate MPEG source files on the main server, making use of the associated index files. The edited programme is stored in its final form on a *finished programme server* ready for use by playout/continuity. As an alternative to a real-time edit conformer, this process could be done by software running in non-real time.
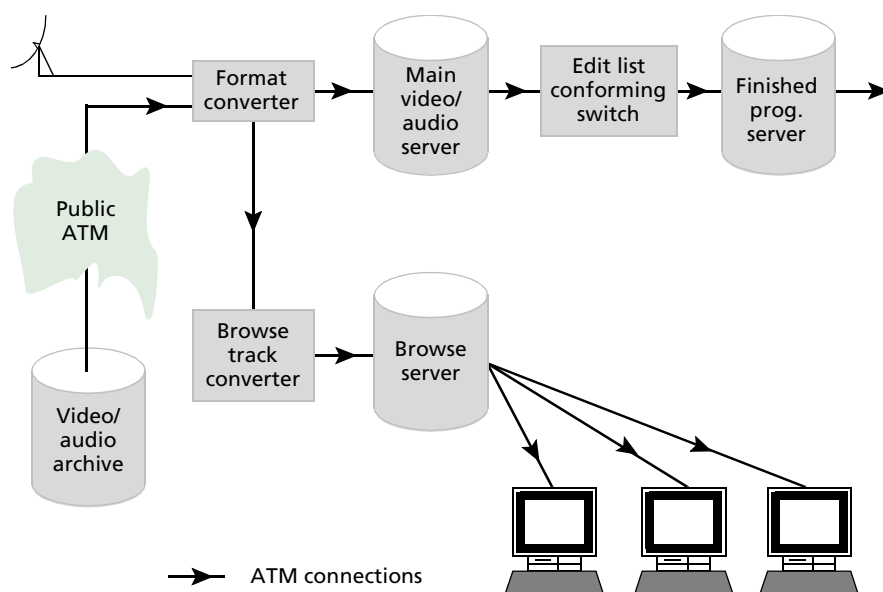
5.2.2.   Network infrastructure

In the ATLANTIC studio reference model of *Fig. 3*, all the functional components are connected together via an ATM network. ATM was chosen for its unique characteristics of flexibility, scalability, provision of bandwidth-on-demand and the ability to support a wide range of quality-of-service requirements (i.e. guaranteed bit-rate) [10].

Within a studio, it is essential to have reliable error-free transmission of data. To meet this requirement it was decided to use the TCP protocol for data transfer since TCP allows for retransmission of any data packets that contain errors. The method chosen for addressing and routeing the data between devices on the network is *Classical IP over ATM*. The performance of such connections has been tested between a range of different platforms and operating systems, and data transfer rates typically in excess of 70 Mbit/s can be maintained over a single ATM connection.

Control of the servers is achieved using protocols which conform to the DSM-CC standard (IS0/IEC-13818-6: Digital Storage Media Command and Control) which is part of the MPEG-2 "family" of standards.

5.2.3.   Decoder synchronization

Within the studio environment there is usually a requirement for a decoder to be synchronized to a studio reference signal. Also, for automatic playout control and for real-time conforming of edit lists, precise control is required of the time that a given decoded frame is displayed at the output of a decoder. Within ATLANTIC this control is achieved by re-stamping all the timing control information within the MPEG bitstream as it passes through the interface from the ATM network to the decoder. This requires that the decoder ATM interface is fed with both SMPTE timecode and the appropriate playout control information in the form of VTR controls or "Louth" server control commands.

6.   Summary

The ATLANTIC project has developed techniques for switching and editing MPEG bitstreams based on transparent, successive, decoding and re-coding of the compressed bitstreams. The techniques involve the use of a MOLE which conveys information about the original video and audio coder decisions within the respective decoded signals.

MOLE-based architectures allow MPEG to be used in a consistent and conventional way throughout all stages of the programme production and distribution chain. Use of MPEG can offer big savings in server sizes, server bandwidths and network bandwidths compared with the use of other compression formats for which the bit-rate is several times higher. These savings could be particularly important for HDTV systems. Also, MOLE-based architectures allow MPEG to be used without loss alongside other alternative compression formats.

Proposals have been submitted to the EBU/ETSI and the SMPTE for standardization of the MOLE signals.

The ATLANTIC project is developing equipment for demonstrations in 1998 of a complete programme production and distribution chain.

Acknowledgements

Bibliography

[1] ATLANTIC Web site: http://www.bbc.co.uk/atlantic

[2] T. Sikora: **MPEG-4 and Beyond – When Can I Watch Soccer on ISDN**
Proceedings of the 20th International Television Symposium, Montreux, June 1997.

[3] M.J. Knee and N.D. Wells: **Seamless Concatenation – A 21st Century Dream**
Proceedings of the 20th International Television Symposium, Montreux, June 1997.

[4] P.J. Brightwell, S.J. Dancer and M.J. Knee: **Flexible switching and editing of MPEG-2 video bitstreams**
International Broadcast Convention (IBC97), Amsterdam, 12-16 September 1997 IEE Conference Publication.

[5] SMPTE standard for Television, as proposed by Snell & Wilcox and the BBC: **MOLE – MPEG Coding Information Representation in 4:2:2 Digital Interfaces.**
ATLANTIC Web site: http://www.bbc.co.uk/atlantic

[6] BBC proposal for SMPTE standard: **Audio MOLE: Coder control data to be embedded in decoded audio pcm.**
ATLANTIC Web site: http://www.bbc.co.uk/atlantic

[7] O.H. Werner: **Generic Quantizer for Transcoding of Hybrid Video**
Proceedings of the 1997 Picture Coding Symposium, Berlin, 10-12 September.

[8] O.H. Werner: **Transcoding of MPEG-2 Intra Frames**
Paper to be published by the IEEE Trans. on Comm.

Nick Wells *graduated from Cambridge University and received a doctorate from Sussex University for studies of radio wave propagation in conducting gases. He has been employed by the BBC at their Research and Development Department since 1977, working mainly in the field of digital video coding for applications within the broadcast chain.*

Dr Wells has actively participated in many standardization activities related to digital TV compression within the EBU, ITU-T and more recently with the ISO/MPEG group. He has also participated in several European collaborative projects such as Eureka 95 for HDTV, the Eureka VADIS Project which co-ordinated the European input to MPEG-2, the RACE HIVITS project concerned with coding for TV and HDTV and, more recently, the ACTS COUGAR and ACTS ATLANTIC Projects.

Nick Wells is currently Project Manager for the ACTS ATLANTIC Project.

[9]   P.N. Tudor and O.H. Werner: **Real-time transcoding of MPEG-2 video bitstreams**
         Proceedings of the International Broadcasting Convention (IBC97) Amsterdam, 12-16 September 1997 IEE Publication.

[10]   A. Alves et al.: **The ATLANTIC news studio: Reference Model and field trial**
         Proceedings of the European Conference on Multimedia Applications Services and Techniques (ECMAST), Milan, 21-23 May 1997.

The transparency of video transcoding

In the accompanying figure, the main processing paths are shown in simplified form for a first MPEG-2 encoder (Coder 1), followed by a decoder and finally followed by a second MPEG-2 coder (Coder 2).

In Coder 1, the difference (a_1) between the source signal and a motion-compensated prediction (mcp_1) is transformed using the discrete cosine transform (DCT). The transform coefficients (b_1) are quantized and coded using a variable-length coder (VLC). The motion-compensated prediction (mcp_1) is formed from previously-coded (and decoded) frames such that the coder and decoder are able to form the same prediction signals.
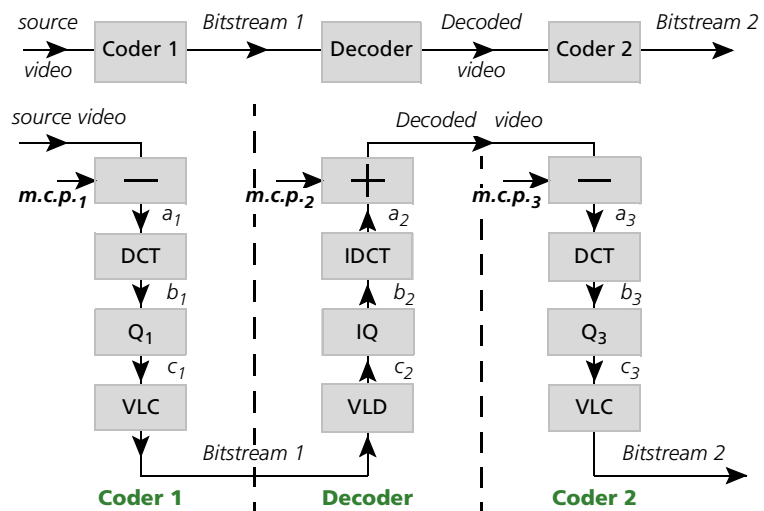

Illustration of transparent coding/decoding/re-coding.

The decoding process is the inverse of this chain. The variable-length decoder (VLD) undoes the variable-length coding; i.e. $c_2 = c_1$.

At its output, the inverse quantizer (IQ) gives quantized coefficient values (b_2) which are fed to the inverse DCT (IDCT). The output of the IDCT process (a_2) is added to a motion-compensated prediction (mcp_2) to give the decoded output. Since, in a standard MPEG-2 encoder, mcp_1 is constructed to be equal to mcp_2, the decoded output is equal to the source signal with the addition of quantization distortion introduced by the combined process of quantization followed by inverse quantization.

The decoded signal is fed into Coder 2. As in Coder 1, a difference is constructed between the input and a motion-compensated prediction, mcp_3. If this prediction can be made equal to mcp_2; i.e. if $mcp_3 = mcp_2$, then $a_3 = a_2$.

(For an I-frame, the prediction is effectively set to zero and therefore $mcp_3 = mcp_2 = 0$ for this frame. Then it can be shown that the predictions of subsequent frames, mcp_2 and mcp_3, derived from this I-frame will be the same provided that the motion vectors and the prediction decisions are identical.)

Since an IDCT process followed by a DCT process is transparent (one inverts the other), then $b_3 = b_2$.

Since b_3 consists of quantized coefficient values, the quantization process Q_3 will not add any further quantization distortion, provided that $Q_3 = Q_1$. Then the process of inverse quantization (IQ) followed by quantization (Q_3) will be transparent, giving $c_3 = c_2$, and therefore, $c_3 = c_1$.

Therefore, $bitstream_2 = bitstream_1$, provided that the second encoder can match the prediction and coding decisions taken by the first encoder. This is achieved through the MOLE.