SEAMLESS CONCATENATION - 
A 21ST CENTURY DREAM

M J Knee (Snell & Wilcox) and N D Wells (BBC)

INTRODUCTION

This paper is about 'seamless concatenation' in the world of compressed television signals, especially MPEG-2 signals, in the context of the session title 'How tightly can we squeeze?'. The paper title has a deliberate ambiguity. It can mean the seamless concatenation of MPEG-2 bitstreams to create a splice or edit, for example to insert an advertisement into an existing bitstream. It can also mean seamless or 'lossless' concatenation of compression coding and decoding operations in a complex broadcast processing chain. Both are a 'dream' in the sense that it would be desirable to perform seamless concatenation, while most embodiments of the MPEG-2 standard fail to make it possible without severely restricting flexibility or compromising on the bitrate savings available from efficient compression.

Why do we wish to perform these concatenation operations on compressed television signals? Let us begin by looking at the kinds of processing that are performed on conventional television signals. These might include:

- real-time cuts, cross-fades, wipes or more complex DVE operations between two sources
- insertion of captions or logos into the picture
- non-linear, frame-accurate editing

As MPEG-2 becomes more widespread as a television format, rather than simply a one-off compression algorithm, it seems reasonable to expect that the same range of functions could be performed on MPEG-2 signals. Compression brings additional requirements, too, because there is no such thing as an 'MPEG-2 bitstream' - there are different bit-rates, levels and profiles, as well as different kinds of bitstream: elementary, PES, program and transport streams. So we need to add to the above list:

- transcoding between different bit-rates, levels and profiles
- transmultiplexing - creating new transport streams from elements of old ones

This paper describes work being done in the ACTS ATLANTIC project to solve the problems of processing MPEG-2 bitstreams in the context of a complete broadcast chain. The techniques described are being made public for the first time at this Symposium. The paper concentrates on video processing because it presents the most difficult challenges and because the benefits of retaining efficient compression are the greatest. However, the ATLANTIC project is dealing with the complete package, including multichannel audio, programme-related and other data and the problems of synchronization between them. General information about the ATLANTIC project is available on a Web site [1].

LIMITATIONS OF EXISTING APPROACHES
The MPEG-2 International Standard and subsequent developments such as the DVB standards have concentrated on getting compressed signals from the studio to the viewer but have paid less attention to what happens to those signals in the complex broadcast networks in between. For example, real-time switching is known to create special difficulties in the MPEG-2 domain. The solutions to these difficulties which have been proposed so far can be classified into three categories:

- Naive cascading
- Restricted MPEG-2
- Bitstream splicing

**Naive cascading**

This approach is where the MPEG-2 signal is fully decoded prior to processing and fully re-encoded afterwards, as shown in Fig. 1.

![Fig. 1 - Naive cascading.](image)

The intermediate processing might be a continuity switch or other effect, or it might be nothing at all but with the re-encoding carried out at a different bit rate or with a different flavour of MPEG-2, constituting a transcoder. In the first case, there is often in a sense no alternative to decoding the picture fully, because most processes carried out on pictures require access to the pixels themselves. However, for the vast majority of the time, a continuity switch does nothing but pass the input picture to the output and we are simply cascading an MPEG-2 decoding and recoding process. In a broadcast chain of even mild complexity, there are likely to be several of these cascades. Research carried out in the ATLANTIC project [2] indicates that there is a significant loss of picture quality associated with multiple cascades, a loss of 5dB being quite typical. This loss can be reduced by using milder compression than is typical with MPEG-2, but in many applications digital bandwidth is an extremely cost-sensitive resource. Remember that we are asking ourselves 'How tightly can we squeeze?'

Such 'naive' cascading is therefore not a real solution to the concatenation problem, though we shall see that the ATLANTIC approach involves what might be called 'intelligent' cascading.

**Restricted MPEG-2**

This approach attempts to avoid the concatenation problem by restricting the compression coding to a limited subset of MPEG-2 so that something approaching frame-accurate editing can be performed, as shown in Fig. 2.

![Fig. 2 - Restricted MPEG-2.](image)
An example is Sony's SX system, where the commonly used IBBPBBP... GOP structure is replaced by an IBIBIB... structure which allows simple edits to be performed largely to the bitstream itself with a minimal amount of intermediate processing of information around the edit point. Such a GOP structure of course requires a higher bit rate for a given level of quality than the IBBPBBP... structure. This is provided for by the MPEG-2 4:2:2 Profile which allows bit rates above the normal limit of 15Mbit/s to be used and which, as its name suggests, also extends the potential vertical chrominance resolution to that of ITU-R Rec. 601. The use of higher bit rates incidentally allows 'naive' cascading to be performed with reduced loss. Other examples with a similar philosophy, though with no claim to MPEG-2 compatibility, are DVCPRO and editing systems based on JPEG compression.

The main benefit of such an approach is that it does offer a genuine solution to the concatenation problem. Unfortunately, this is at the expense of incompatibility with mainstream MPEG-2 coding. In a closed environment this might be acceptable, but the very success of MPEG-2 Main Profile means that it will increasingly be used as the medium for contribution (e.g. from SNG equipment), primary and secondary distribution (e.g. in digital terrestrial television systems) and archive storage (e.g. in the ACTS AURORA project). In such a world, the use of a restricted version of MPEG-2 for editing purposes, and the consequent need for transcoding between Main and 4:2:2 profiles, becomes a much less attractive option. The need for a fairly high bit rate also undermines the benefits of using compression in the first place.

**Bitstream Splicing**

This approach is an attempt to solve the problem of simple bitstream switching, for example for ad insertion. It does not work for any of the more complex functions such as cross-fades or caption insertion, and it does not address the quality issues of transcoding.

The SMPTE has carried out significant work in this area. Its Working Group PT20.02 (Switching and Synchronization) set up an ad-hoc group that has proposed a standard for bitstream splicing [3]. Two types of splice are proposed: 'seamless' and 'non-seamless', depending on how a decoder behaves when it receives the spliced bitstream. Seamless splicing is suitable for certain types of switching but imposes restrictions on flexibility as to when the switch can be performed. These can be partially overcome by manipulating the upstream coder's buffer management system to provide several splice points, but this is likely to lead to some loss of quality. Non-seamless splicing could also be used, but still lacks some flexibility and imposes requirements on downstream decoder behaviour. For good results, programmes should end with a black frame.

The disadvantages of this kind of bitstream splicing may be summed up by the observations that

- it lacks flexibility, both in choice of splice point and in the kind of switching that can be performed
- upstream coders need to know where bitstream splicing might be performed
- downstream decoders need to behave in a certain way, particularly where buffer underflow is concerned

The ATLANTIC approach has none of these disadvantages. However, the bitstream splicing approach might still be useful if a low-cost solution were required for applications where the restrictions are not a problem.

**THE ATLANTIC APPROACH**
The Information Bus

The ATLANTIC approach to handling MPEG-2 bitstreams recognizes that most processing operations - in fact, anything beyond the simple kind of splicing discussed above - require access to decoded pictures. There is therefore an inevitable cascade of decoding and re-encoding operations. What differentiates the ATLANTIC approach from 'naive' cascading is the following:

- in naive cascading, 'side information' (e.g. motion vectors and coding mode decisions) decoded from the bitstream is used to generate the decoded picture and is then thrown away. The re-encoder has to re-make all these decisions.
- in ATLANTIC cascading, the side information is transferred to the re-encoder and is used in its decision processing

In the ATLANTIC approach, the side information is given a formal data structure and is known as the 'Information Bus'. A simple cascade using the Information Bus is illustrated in Fig. 3.

![Fig. 3 - An ATLANTIC cascade.](image)

The ATLANTIC Decoder is a standard MPEG-2 decoder but with an additional output, the Information Bus, which is synchronized to the video output. The Dim Coder is the core of an MPEG-2 encoder but takes all its coding decisions from the Information Bus, ensuring that the Dim Coder takes the same decisions as were taken by the upstream coder. Indeed, it is possible to demonstrate mathematically that the only cascading impairments introduced by this ATLANTIC decoding and re-encoding process are the tiny ones due to mismatches between the DCT and inverse DCT functions. In practice, these have been shown to amount to about 0.0002dB, clearly a negligible loss.

Of course, merely decoding and re-encoding the bitstream as a permanent transparent process would be useless in itself. The benefits in cascading using the Information Bus only become apparent when the simple cascade is an instance of something more flexible. We now go on to describe the two main kinds of process that use the Information Bus. These are:

- picture operations, in which both the picture and the Information Bus are processed before re-encoding takes place
- transcoding, in which no picture processing takes place but the re-encoding is performed with some parameter changes, in particular a change in bit-rate

Picture processing using the Information Bus

A bitstream switch
The first kind of process will be illustrated using a simple example: a bitstream switch. While we have seen that this particular operation can sometimes be done using splicing techniques, the Information Bus technique allows a completely unrestricted switch to be performed. The extension of the technique to cross-fades and wipes will also become apparent.

A block diagram of a bitstream switch using the Information Bus is shown in Fig. 4.
In the steady state, when bitstream A is selected, the Information Bus processor is passing the Information Bus from ATLANTIC Decoder A unchanged to the Dim Coder, so we have a simple cascade of the kind shown in Fig. 3, so the output bitstream is essentially identical to input bitstream A. The same situation occurs in the steady state after the switching operation, when bitstream B is selected.

The switching operation itself is carried out on the decoded video signals, so these have to be synchronized. This can be done in the bitstream domain by extending the ATLANTIC decoder buffers so that additional delay can be inserted into the bitstream, or it can be done in the decoded video domain using conventional synchronization techniques. Near the switching point, neither Information Bus A nor Information Bus B is directly usable as the source of re-encoding decisions because reference will be made to picture information on the wrong side of the switching point. For example, if the picture types (in display order) and switch point are as shown below, the Information Bus will only be valid during the periods shown in bold type:

**Sequence A:**


**Sequence B:**


Switch point: *

Output sequence: **I B B P B B P B B P ?? ?? ?? ?? ?? I B P B B P B B P**

During this 'switching period' the re-encoder is 'on its own'. It has to work as a full MPEG encoder would work on a scene change. However, it can make use of the Information Bus signals of the two bitstreams to aid its coding decisions. The elimination of cascading impairments made possible by the Information Bus is no longer possible during this period, but this is a fundamental limitation because the picture to be coded is new information from an MPEG point of view.

It is the goal of the re-encoder to work in such a way that, as soon as possible after the switch point, the re-encoder can be locked to the new Information Bus so that cascading impairments are eliminated. While most coding decisions can be locked in this way as soon as the switching period is over, there may need to be an additional 'recovery period' during which the quantizer in the re-encoder is adjusted until the encoder buffer status is correctly related to the vbv\_delay value in the Information Bus. During this recovery period, the re-encoder is actually working as a transcoder using the techniques described below. When this convergence is reached, the re-encoder can be locked to the Information Bus and the steady state 'ATLANTIC' cascade is once again attained. In this situation, it would actually be possible (though unnecessary) to switch the decoder-recoder combination out of circuit altogether and replace it seamlessly by a compensating delay. This demonstrates that in fact the only calculations required to perform the switch take place near the switching period. In an editing system using the ATLANTIC switch, it would therefore be possible to implement
the switch efficiently using non-real-time software processing.

One point to note from the above example is that there is no restriction in the relative or absolute GOP structures of the two bitstreams. In the example, bitstream A has a regular, common, IBBPBBP... structure but bitstream B has an almost random-looking but perfectly legal structure. Such arbitrary parameters in the input bitstreams do not present a particular problem for the ATLANTIC switch.

Further details on how coding decisions are made during the switch period, and on how the recovery period is managed, can be found in [2].

**Introducing the Mole**

We can now see that, in principle, the switch could be replaced by a vision mixer or DVE machine, enabling a full range of functions to be performed on the decoded pictures. The only problem with such a generalization of the ATLANTIC switch is in the handling of the Information Bus signals, which would have to be passed round the mixer and processed separately, making use of information about the status of the mixer.

This remaining problem is overcome by converting the Information Bus into a special format known as the 'Mole' which enables conventional digital studio mixing or DVE equipment to be used. Whenever the studio equipment is passing one of its input signals untouched, the Mole associated with that signal is also passed automatically, 'burrowing through' the studio equipment to emerge at the other end ready for use in re-encoding. Conversely, whenever the equipment is affecting the signal in any way, for example during a cross-fade, the Mole is automatically destroyed and cannot be used for re-encoding. Fig. 5 shows an ATLANTIC switch using the Mole.

![Fig. 5 - ATLANTIC switch using the Mole.](image)

The Information Bus is inserted into the video signal by the 'Mole composer' at the output of each ATLANTIC decoder and is decoded by the 'Mole interpreter' at the input to the ATLANTIC encoder.

In order for the Mole to work as a means of passing the Information Bus through studio equipment, several conditions must be satisfied:

- The Mole must be invisible on the video signal
- The studio equipment must be capable of passing the whole digital signal untouched, without any hidden filtering, noise reduction or other processing
- The ATLANTIC encoder must be capable of detecting when the Mole is no longer valid so that it knows when it has to make its own coding decisions

The requirement that the studio equipment should be transparent to the digital signal seems a
strict one, but it should be remembered that, as a general rule, if a signal is changed it becomes a new source as far as the re-encoder is concerned and needs to be coded afresh.

It is important to note that the ATLANTIC technology imposes no special constraints either on upstream encoders or on downstream decoders, which can be standard 'off-the-shelf' equipment. However, it is clearly desirable that the quality of the first compression encoder in the chain is as high as possible, because this will define the quality throughout the chain.

Prototype ATLANTIC decoders and encoders, including the Mole processing outlined above, are currently under construction in the ATLANTIC project.

**Transcoding**

In this section we shall discuss how bit-rate transcoding is performed using ATLANTIC technology. The basic principle of passing the Information Bus from a decoder to a Dim Coder is used, but this time there is no need for intermediate pixel-based processing and hence no need for the Mole.

In the ATLANTIC project, two kinds of transcoder are being investigated. The first is a 'drifty' transcoder, illustrated in Fig. 6.

![Fig. 6 - A 'drifty' transcoder using the Information Bus.](image)

Here, the decoding pipeline stops at the generation of inverse-quantized DCT coefficients. These coefficients are then requantized under the control of a bit-rate control algorithm set to the new bit rate. The Information Bus is simply passed from the variable-length decoder to the variable-length encoder for insertion into the new bitstream. This kind of transcoder is called 'drifty' because the predictions generated in the downstream decoder will not match those generated in the upstream encoder, causing errors to be accumulated in successive P pictures through the GOP. Another disadvantage with this architecture is that there are few opportunities for changing coding parameters such as prediction modes. In some circumstances, for example, transcoding to a slightly lower bit rate, these restrictions may be acceptable and may be considered a price worth paying for a very simple transcoder. A detailed discussion of drift can be found in [4].

The second kind of transcoder is a 'full' transcoder in which the decoder and encoder prediction loops are implemented, as shown in Fig. 7.
Even here, some decoding and re-encoding steps can usually be left out, such as 4:2:0 to 4:2:2 conversion and picture re-ordering. With this architecture, no drift is introduced and it is possible to change some of the coding parameters. However, if the parameters related to prediction are left unchanged, it is possible to simplify the architecture so that it contains only one prediction generator, as described for example in [5].

In both kinds of transcoder, the requantization step is a very important one. In all MPEG encoding, including the re-encoding in a transcoder, the inverse quantizer reconstruction levels are specified in the standard. However, the quantizer decision levels are not specified and can be chosen to optimize picture quality at a given bit rate. In the case of transcoding, it is possible to choose decision levels on the basis of knowledge of the previous quantization process. Work in the ATLANTIC project [6] has shown that such a requantizer can give significantly better picture quality at a given bit rate than 'naive' use of a commonly used quantizer such as that defined in the MPEG Test Model.

This brief discussion has concentrated on a transcoder as a stand-alone unit. However, the 'full transcoder' approach can equally be applied to any switching or other processing equipment using the Mole. In fact, transcoding is necessarily carried out during the 'recovery period' following a switch, and the architecture makes it possible to specify an arbitrary bit rate at the switch output or to deliver arbitrary bit rates at the inputs so that transcoding is automatically carried out whenever necessary.

In the discussions of the ATLANTIC and other approaches at the beginning of this paper, we stressed the fact that the ATLANTIC approach will handle any MPEG-2 signals without restriction. This is certainly true of transcoding, but a possible drawback should be pointed out. This is that bad coding decisions and motion vectors cannot be made better by ATLANTIC technology alone. If a relatively high bit-rate bitstream that does not make proper use of the MPEG-2 specification (e.g. has a very small motion vector range) is received for transcoding to a low bit-rate, the re-encoder will essentially have to work as if it were encoding for the first time. We can conceive of a range of transcoders with varying degrees of capability for remaking coding decisions, but the best approach is to generate good quality bitstreams the first time they are encoded. High-quality source coding, including pre-processing, was the subject of the COUGAR project, whose very successful results are
reported in [7] and [8].

A prototype transcoder implementing both the approaches outlined above is currently under construction in the ATLANTIC project.

**FURTHER EXAMPLES OF THE USE OF ATLANTIC TECHNOLOGY**

**MPEG editing**

The MPEG switching technology described in the previous section makes it possible to develop a frame-accurate, nonlinear editor in which the input and output interfaces and the information stored on the server are all in MPEG-2 form. Such an editor has two major advantages over other technologies:

- Editing can be performed in the MPEG domain without fear of introducing artefacts due to cascading
- MPEG-2 is much more efficient in bit rate terms than other formats usually used in editors, leading to greatly reduced costs for servers and for network connections

Fig. 8 gives a block diagram of a post-production system, suitable for small studio applications, which is being built in the ATLANTIC project.

![Block diagram of the ATLANTIC post-production system.](image)

In this ATLANTIC system, video and audio sequences are transferred via an ATM network using standard TCP/IP file transfer protocols, although any other suitable network and protocol could be used.

The following is a brief description of the main elements of the editor. Further information can be found in [9].

The **Format converter** converts incoming MPEG-2 transport streams into separate video, audio and data PES streams for storage on the main server, and generates an index file.

The **Browse track generator** generates an MPEG I-frame-only, reduced resolution browse
track corresponding to the video stored on the server, and possibly an audio browse track, and stores these on the browse server.

The **Journalist Workstation** operates as a conventional edit workstation on the I-frame browse track, providing a graphical user interface to allow the journalist to create an edit decision list.

The **Edit Conformer** is in fact the real-time ATLANTIC switch, controlled by the edit decision list to pull bitstreams from the main server, conform the edit and place the results onto the finished programme server.

The development of a true MPEG-based editor poses several design challenges which the ATLANTIC project is addressing. These include:

- audio/video synchronization
- handling of programme-related data
- audio editing itself; initially, a simple bitstream switch operating on audio frame boundaries will be used, but for more refined operation an audio 'Mole' has been developed to enable more sophisticated processing to be carried out with the same benefits in reduction of cascading impairments as the video Mole provides
- the possibility of replacing the (initially rather expensive) edit conformer with a software version
- design of a low-cost journalist workstation using off-the-shelf MPEG decoders
- control of the whole system

While the challenges are serious, the potential advantages of an efficient MPEG-based editor are enormous. The ATLANTIC approach has indeed made the '21st-century dream' a reality.

**Dynamic transmultiplexing**

Another 'dream' related to the manipulation of MPEG bitstreams is the possibility of a 'dynamic transmultiplexer'. This is essentially a remultiplexer incorporating transcoders, giving it the capability to change the video bit rates of any or all of the services that will go to make up the output transport stream. The transcoders may be controlled either automatically by the remultiplexer, allocating bit-rate on a dynamic basis, or manually by the user. In either case, they make use of the ATLANTIC technology described above to optimize the transcoded picture quality. The output buffers of the transcoders have to be controlled in such a way that changes in bit rate do not cause any buffer violations or discontinuities in the downstream decoders. Details of how this is achieved are given in [10].

The dynamic transmultiplexer using ATLANTIC technology is a key component in complex digital television networks, for example digital terrestrial networks with regional variations.

**CONCLUSIONS**

This paper has described technology that makes possible seamless, unrestricted switching, mixing and other operations, together with high-quality transcoding, on any MPEG-2 bitstreams. The technology is based around the 'Information Bus', the 'Mole' and the optimization of transcoding quality. Explanations of the technology have been given based on a bitstream switch and a transcoder. Two more sophisticated examples of use of the technology have been described: an MPEG-based editor and a dynamic transmultiplexer. Between them, these components make possible a full range of processes on MPEG bitstreams while preserving quality throughout a complex broadcast chain, avoiding the need to use high bit rates to avoid cascading impairments.
In summary, the answer to the title of the session in which this paper is being presented: 'How far can we squeeze?' is 'As far as MPEG-2 will allow for the quality desired; there is no need to give up the cost benefits of compression to buy flexibility in bitstream processing'.

ACKNOWLEDGEMENTS

The ATLANTIC project is being supported by the European Commission within the ACTS framework.

The partners in the ATLANTIC project are BBC (UK), Snell & Wilcox (UK), CSELT (Italy), INESC (Portugal), EPFL (Switzerland), ENST (France) and FhG (Germany). The contribution of all these partners to the work described in this paper is gratefully acknowledged.

The authors also wish to thank their respective organizations, Snell & Wilcox Ltd. and the British Broadcasting Corporation, for permission to publish this paper.

REFERENCES

1. ACTS 078 ATLANTIC project Web site: http://www.bbc.co.uk/atlantic/