In this article the Author considers, firstly, a hypothetical broadcast network in which compression equipments have replaced several existing functions – resulting in multiple-cascading. Secondly, he describes a similar network that has been optimized for compression technology.

Picture-quality assessment methods – both conventional and new, subjective and objective – are discussed with the aim of providing background information. Some proposals are put forward for objective evaluation together with initial observations when concatenating (cascading) codecs of similar and different types.

1. Introduction

Bit-rate reduction (BRR) techniques – commonly referred to as compression – are now firmly established in the latest generation of broadcast television equipments.

Early applications of compression in broadcast studio equipments were mainly limited to off-line editing and similar processes using the JPEG compression system. JPEG was originally specified for still-picture applications in the printing industry and was adapted to motion video applications, by encoding on a picture-by-picture basis. Various proprietary versions of JPEG, with higher compression ratios, were introduced to reduce the storage requirements, and alternative systems based on Wavelets and Fractals have been implemented. Due to the availability of integrated hardware, motion JPEG is now widely used for off- and on-line editing, disc servers, slow-motion systems, etc., but there is little standardization and few interfaces at the compressed level.

The MPEG-1 compression standard was aimed at progressively-scanned moving images and has been widely adopted for CD-ROM and similar applications; it has also found limited application for broadcast video. The MPEG-2 system is specifically targeted at the complete gamut of video systems; for standard-definition television, the MP@ML codec – with compression ratios of 20:1 and greater – is now firmly established by the DVB standards for satellite, cable and terrestrial transmissions. Significantly, MPEG-2 specifies the interface data-stream, thus providing interna-
tional standardization of interfaces at compressed levels.

MP@ML is not suited for studio applications, due to the 4:2:0 digital encoding it uses and also through its use of a lengthy frame structure Group of Pictures (GOP), which is not convenient for editing or other frame-accurate studio processes. A professional or studio level has now been introduced, MPEG-2 422P@ML, which has more modest compression ratios of between 5:1 and 15:1, and short GOPs, which allows frame-accurate processing while maintaining a high picture quality with multi-generation capabilities.

The natural progression of development and introduction of new technologies will, in time, encourage the rationalization of compression systems to a few types, probably based on the family of MPEG-2 MP@ML and 422P@ML specifications. In the meantime, broadcasters are faced with a large selection of different equipments, promoted by different vendors all claiming particular operational, cost-saving and performance advantages or improvements over the more conventional uncompressed techniques.

How can a user decide on how and where to adopt compression systems, and what are the economic, operational and performance implications? How is the picture quality to be measured and can it be predicted in complex systems?

2. Application of compression systems

By way of example, two illustrations will be considered for a broadcast chain from source to transmission. The examples are hypothetical and may appear somewhat complex but could be representative of a large national digital network, comprising several regional production and transmission centres. In the first example, individual processes throughout the chain have been replaced by equivalent systems which incorporate internal compression but interfaced at the standard digital levels of ITU-R Recommendations BT.601 [1] and BT.656 [2]. The second example considers a similar chain, but conceived on an overall integrated strategy.

2.1. Conventional broadcast chain

The complete chain is illustrated in Fig. 1 which has been annotated "C_n" where compression products might replace conventional digital installations. Typical options are outlined below.

It is assumed that the acquisition ENG camera would have an integral digital tape or disc-based recording system which uses compression (C_1). The present preferred format is tape and there are now many format options, ranging from the modest compression rates of professional 4:2:2-based systems to the higher compression rates of systems developed from the consumer VHS and DVC video recorder formats. Most of these tape formats, together with the single disc system to be demonstrated so far, have field edit options. The source video signal, reconstructed to 270 Mbit/s full-rate SDI, must be linked as a contribution to the studio centre. It is unlikely that the full data-rate can be accommodated on this link and, thus, compression (C_2) would have to be used: the options available here range from, for example, ETSI 34, 17 or 8 Mbit/s DCT systems to 8 Mbit/s MPEG-2.

---

**Abbreviations**

422P@ML 4:2:2 Profile at Main Level
ATM Asynchronous transfer mode
BRR Bit-rate reduced
CCIR (ITU) International Radio Consultative Committee
DCT Discrete cosine transform
DVC Digital video cassette
ENG Electronic news gathering
ETSI European Telecommunication Standards Institute
GOP Group of pictures
ISO International Standards Organisation
ITU-R International Telecommunication Union, Radiocommunication Sector
JPEG (ISO) Joint Photographic Experts Group
MOSAIC Methods for Optimization and Subjective Assessment in Image Communications
MPEG (ISO) Moving Picture Experts Group
MP@ML Main Profile at Main Level
PSNR peak signal-to-noise ratio
RMS Root-mean-square
SDDI (Sony) serial digital data interface
SDH Synchronous digital hierarchy
SDI Serial digital interface
SMPTE (US) Society of Motion Picture and Television Engineers
The received decoded contribution, now back at 270 Mbit/s, may require temporary storage, possibly within a video server. Although uncompressed working is possible, economic considerations and maximum storage capacity will demand the use of compression (C3) which most likely will be based on motion JPEG. Further editing with other contributions might be required, using a workstation with integral disc storage, also based on motion JPEG (C4), but communicating to and from the temporary storage unit at full-rate 270 Mbit/s SDI.

Within the studio centre, communications and studio production processing are likely to use full-rate SDI, since many standard production operations cannot function with compressed signals. However, within the production process, JPEG compression (C5) might be present for slow-motion effects or similar, and material to and from an archive may be subject to digital VTR compression (C6).

Selection of the produced programmes for network distribution is controlled within a presentation area which may also insert commercials, promotions, and network idents from a presentation video server, again using JPEG compression (C7). This server may also function as a temporary store for pre-recorded programmes.

The final output from the studio centre is digitally distributed to the regional centres which, for economy of bandwidth, might use compression at 140 or 34 Mbit/s (C8), depending on the distribution medium available, i.e. fibre or satellite.

At the regional centre, the decoded full-rate SDI signal will require the insertion of local regional material or commercials from a compression-based presentation server (C9). Alternatively, the signal may be temporarily stored, for time-delay purposes, within the same server or on a digital VTR which uses compression (C10).

The final regional output is then subject to secondary distribution and transmission to the viewer. In so doing, two further compression systems could be involved: the first for communication to the transmitter, possibly using ETSI DCT (C11) and the second for final conversion to the compressed transmission format (C12) which, in all probability, will be MPEG-2 MP@ML.

**Figure 1**
Conventional broadcast chain, incorporating compression products.
The piece-meal incorporation of equipments using compression systems, as illustrated above, is clearly not ideal. In this particular example, there are some twelve different compression systems of which ten or eleven could be concatenated in the complete chain from source to transmission. The most significant disadvantages of this network, in no particular order, are:

(a) The signal quality is reduced due to the concatenation of different types of compression codecs, operating over a wide range of compressed data-rates. Each codec contributes its own approximations or noise error signals, which can accumulate with concatenation. The total error will be dominated by the codec with the greatest compression but this can be enhanced by earlier coding errors. Compression coding failure could occur with certain types of picture material.

(b) With a large number of different compression systems and equipment vendors, there are no economies in capital equipment purchases.

(c) The number of different compression systems place extra demands on operational staff or on the protocol requirements of an automation system.

(d) Training of maintenance staff is increased due to the number of different systems.

The only apparent advantage of the piece-meal approach is that capital expenditure can be spread over a long period.

With regard to the overall picture quality of this example broadcast chain, it is clear that some form of objective quality rating for individual codecs would be desirable, together with rules for the accumulation of ratings when a number of codecs are concatenated. Before embarking on a discussion of quality assessment, it is worth considering a possible alternative to the above “compression islands in a digital sea” scenario, which will reduce the number of concatenated codecs in the broadcast chain.

## 2.2. Integrated broadcast chain using compression technology

Fig. 2 is a conceptual diagram for a complete broadcast chain, from source to transmission, which uses compression technology, and digital interfaces for the transfer of data in a compressed format.

The actual compression format is not relevant here although use of a family such as MPEG-2 422P@ML would be recommended, by reason of its standardization and the expected operations that should be possible in the compressed domain. For final transmission, and for contribution where bandwidth is at a premium, it will be necessary to use MPEG-2 MP@ML compression or similar. Likewise, the interface specification is not important as it is assumed that, within an operational area or a complete studio centre, all equipments operating in the compressed domain would be linked by a studio centre network or bus structure such as Ethernet, SDDI, Fibre Channel etc., using SDH, ATM or a similar transport protocol.

The acquisition ENG camera and the field edit system could incorporate any compression system but preferably would use a standard 422P compression (C₁) which would also be that used throughout the complete chain. Ideally, the contribution link to the studio centre will accommodate the associated 422P data-rate of 20 – 50 Mbit/s. However, the available link may have excess or limited capacity in which case the communication could be faster than real time, or slower than real time. If live contribution is essential with a limited-capacity link, then it will be necessary to compress the signal further (C₂).

At the studio centre, the received signal – restored if necessary to the original C₁ compression rate – could be connected via the studio compressed network to a temporary storage video server. This server will also hold other source material or transferred archives which have been restored or converted to the C₁ rate where necessary. Further editing can be conducted within workstations, probably operating internally with more highly-compressed signals (C₃), but the final conforming would take place within the temporary video server at the main C₁ compression rate.

As with the conventional chain, studio production will need to operate at full-rate SDI (270 Mbit/s) to give the greatest flexibility for studio processing. A possible method of achieving this flexibility is via a second studio centre network or bus, operating at the full SDI rate, with a number of compression codecs transferring data between the compressed network and the SDI network. The number of codec transfers between the two studio networks can be minimized through the availability of a video server, or VTRs, operating with full-rate uncompressed SDI signals and normally assigned to the production studio. The server could also provide slow motion and other production effects that require storage. The routing of signals between the studio networks, via the compression codecs, could be automatic by reason of integral signalling which will identify the format of the routed signal as being either compressed or full-rate SDI. Archive material may be compressed or
full-rate and would also be automatically routed over the appropriate studio network. In the event, it should be possible to limit the number of codec transfers between the studio networks at the C₁ rate to just one or two, thereby minimizing the number of concatenated compression codecs used during studio production.

In a similar manner, the presentation area which selects the programmes and inserts the commercials, promotions and idents for distribution over the national network would work from either the compressed or the full-rate SDI studio networks. The presentation area would also have access to compressed and uncompressed video servers. Simple presentation-switching operations could be effected directly on the compressed signals, but full uncompressed production operations would also be available if more sophisticated presentation is required. Once again, the objective is to reduce the number of compression codec operations. The final output from the studio centre would be distributed digitally to the regional centres at the compressed C₁ rate, which is feasible for most distribution media.

Within the regional centre, the presentation area would be similar to that of the main centre described above. Compressed C₁ and full-rate SDI local networks would both be available, and local regional material or commercials could be switched at compressed level, or at full SDI level.
for more complex presentation. Conversion of the regional output signal from C1 compression to the transmission format could occur at the output of the regional centre or at the transmitter, depending on the capacity of the secondary distribution links and the number of transmitters in the region. It is believed that conversion between MPEG-2 422P@ML and MPEG-2 MP@ML is feasible without decoding to full-rate 270 Mbit/s SDI.

The complete chain just described is purely theoretical although some pilot production centres which operate on similar principles have recently been commissioned. The objective of reducing the number of concatenated codecs has been achieved and, in the worst case, is four or five with the possibility of an additional different format if the source contribution link has insufficient capacity. The advantages of this network are:

a) the number of concatenated codecs is minimized;

b) high picture quality is assured since most codecs are of the MPEG-2 422P@ML type;

c) there is an economy of scale due to use of unified standard compression equipment;

d) the operations are integrated;

e) maintenance is simpler and rationalized.

A possible disadvantage of the integrated approach is that, in the short term, such a system may only be available via single vendors who are offering proprietary features. This may be considered advantageous to some users but, with the passage of time, these features should become less significant as equipment designs merge towards compatible standards.

### 3. Objective assessment of compression codecs

Currently, the assessment of compression systems is still subjective which, of course, for the broadcast viewer is the main aim. Indeed, many picture defects are hidden by the picture transition or the content with which they are associated. However, the broadcaster requires more information on the manner in which a compression system may have modified the picture, and the subsequent effects of further compression or other broadcast processes [3].

New assessment techniques are evolving but, before discussing some possibilities, is it worth considering the various methodologies in current use.

#### 3.1. Static test waveforms – objective assessment

Traditional methods of evaluating a broadcast video system have used a suite of standard test waveforms and it might be considered that these techniques can be extended to digital compression systems. However, the static nature of these tests will not fully exercise the compression algorithms. Even when the test signals are digitally generated and also digitally interfaced to the system under test, there is little random or quantizing noise to stress the system further.

High-energy frequency-domain signals, such as sweep or multi-burst, may reveal pre-filter characteristics or – in the case of simple compression systems – buffer-overload or coding limitations. Nevertheless, a complete conventional automatic test on a compression system will normally indicate exemplary objective performance, even on systems which reveal subjective shortcomings on normal programme material. Many of the standard test signals were devised to provide an objective figure, based on a subjective assessment, at a time when distortions were limited to analogue linear and non-linear parameters – the temporal domain being of little significance.

#### 3.2. Moving test sequences – subjective assessment

Assessment of early BRR systems gave rise to the concept of subjective evaluation using moving test sequences. Inevitably there were many sequences preferred by various proponents, mainly extracted from standard programme material to highlight, or maybe to hide, certain system attributes.

In an attempt to standardize subjective evaluation, a library of sequences was developed jointly by the EBU and CCIR (now ITU), suitable for the testing of compression systems developed for use within the standard telecommunication hierarchies of 34 and 45 Mbit/s. These libraries have achieved wide circulation and many of the specially-devised – but nevertheless, realistic – sequences are widely used. Typically common are “Mobile and Calendar” and “Renata with Butterflies”, both of which contain high-saturation moving detail, along with realistic camera attributes of noise and lag. A further but atypical sequence “Diva with Noise”, which was specifically devised to be extremely testing, is also frequently used to reveal system deficiencies. The library also contains some simpler sequences such as “Susie”, a portrait of a lady using a telephone, which should be transferred through a compression codec with little if any distortion.
Assessment of these moving sequences is by subjective viewing, using one of the standardized methodologies. Such evaluation is time-consuming, requiring special viewing conditions, a large number of observers and, preferably, independent evaluations in more than one assessment centre. There has been some progress in the development of viewing models [4][5][6][7] which have a fair correlation with subjective results, but the Author feels that it may be some time before such a model can be simplified for universal acceptance and application. Quality issues are not confined to the studio; the viewer’s tolerance to distortion is also important. This issue is being addressed by a European Commission RACE project, MOSAIC, which is developing reliable methods for subjectively judging the picture quality of long-duration picture sequences, using untrained observers (see the article commencing on page 12 [8]).

An alternative, synthesized, moving sequence [9] – available now from a digital generator and also from a D1-format videotape – is designed to test the limits of the JPEG, MPEG-1 and MPEG-2 compression codecs. The sequence contains components for analyzing the probable algorithm content of a compression system and its comparative performance. By reason of the artificial content, it is not a sequence suitable for subjective testing but may offer possibilities for objective evaluation.

The well known zone plate pattern can have temporal components and has been used for codec evaluation to reveal certain properties by subjective evaluation. Once again, it is not suitable for full subjective testing and there are no standard procedures for objective evaluation. However, the zone plate pattern offers possibilities for new analytical techniques.

### 3.3. Moving test sequences – objective assessment

The time-consuming nature of subjective testing, allied with the fact that the test principle is based only on visible artefacts, implies that the broadcaster still requires some form of objective evaluation of codec performance. There are two requirements for an objective rating:

1. A rating which correlates with subjective evaluations, so that the broadcaster can forecast the viewer’s acceptance of compression artefacts.
2. An objective rating for the absolute value of compression-coding errors, and rules for accumulation when codecs are concatenated. With

![Diagram](image-url)
such a tool, the broadcaster can ascertain that a defined level of picture quality can be maintained throughout a broadcast chain and can predict the level of concatenation at which failure may occur.

A possible technique for objective evaluation is one of examining the activity or entropy of the difference between the predicted and the instantaneous signals within the feedback path of a compression encoder. This technique has been used successfully to confirm the criticality of test sequences, but relies on internal access to the encoder which is not always possible. It is possible, however, to subtract the system input and output signals which, by reason of the general architecture of compression codecs, is similar to the encoder prediction signal (see Fig. 3). The input/output difference signal is the error signal noise and is defined as the Peak Signal to Noise Ratio (PSNR). A codec with low errors has a high PSNR. The error signal noise is bipolar and has similar properties to more conventional video noise, the amplitude of the error being dependent on the “busyness” or entropy of the signal. A digitally-generated static test signal has negligible or no error noise, whereas a moving sequence has noise which depends on picture content.

Computer graphic sequences tend to exhibit noise at transitions or changing areas, but large noise-free areas have negligible errors. Camera-sourced pictures, with low levels of residual random noise, exhibit an overall compression error noise together with picture content contributions. In the limit, extreme test sequences such as “Diva with Noise” exhibit high error noise components.

An interesting procedure for analyzing the modification is to highlight the output picture with false colours, or similar pixel-sized markers, where errors exceed a pre-determined level. This confirms that most modification occurs at transitions or areas of high activity, due to picture content or noise. The “noise” nature of the difference signal allows objective analysis by means of noise measurement and spectral analysis, using standard instrumentation. The noise measurement gives an absolute error value whereas the spectral analysis can give information on the nature of the errors, i.e. transitional, blocking etc. which, in turn, could determine a form of weighting function for the noise measurement.

The choice of picture sequence and the exact method of quantifying the noise measurement is still under investigation. For the measurements reported here, a 10-second segment of “Mobile and Calendar” was used – being typical of a real camera source with moderate detail and movement – and the noise was measured as the mean RMS value over the duration of the sequence.

4. Concatenation of compression codecs

A number of error-signal noise measurements have been conducted on single and concatenated codecs, mainly on different realizations of the ETSI 34 Mbit/s video codec [10], as it is regarded as typical for codecs based on DCT and temporal

![Figure 4 Cumulative error noise for concatenated ETSI 34 Mbit/s codecs.](image)
prediction. Codecs for MPEG-2 MP@ML were not generally available at the time of the investigation.

When codecs are concatenated, the overall PSNR is a function of several factors, the most important being:

a) The PSNR generally accumulates according to a root-mean-square law (see Fig. 4). Thus for two similar codecs, the PSNR is reduced by approximately 3dB. For dissimilar codecs, the PSNR of the higher compression unit dominates but errors from the lower compression unit can still reduce the overall PSNR.

b) For codecs which are completely identical in all coding parameters (including frame or GOP synchronization), the PSNR is determined by the first codec only, except if there are picture alterations (processing) between codecs which modify the level, spatial or temporal position (shift, or picture size and shape) etc., in which case the overall PSNR accumulates by the root-mean-square law.

c) A high error-rate codec, with a low PSNR, may cause a following codec to limit in the redundancy removal process, resulting in visibility of the processed picture blocks.

A further technique for data reduction is to pre-process or pre-filter the picture signal prior to the main compression algorithm. There can be simple filtering of the colour components – for example, horizontal filtering (4:1:1) and vertical filtering (4:2:0) – or more complex spatial and temporal filtering which could include dropping the alternate fields. Concatenation of different pre-filtering processes can result in a further loss of picture quality.

As mentioned above, the spectrum of the error signal noise can indicate the nature of the errors and this is well illustrated in Fig. 5 which shows the noise spectrum for the first, second and third concatenations. The strong harmonic components, which arise with later generations, indicate blocking due to errors in the low-frequency components of the DCT blocks. Such blocking is subjectively disturbing and thus the spectral presence would indicate that an adverse weighting factor should be applied to the noise measurement.

This summary can only offer an approximate indication of the complex processes which occur when compression codecs are concatenated. Various investigations have been reported [11][12][13] and work continues on techniques for the objective assessment of compression-codec performance [14][15].

For the present, it is important to state, once again, that the piece-meal incorporation of compression systems into an existing broadcast chain should be undertaken with caution and with due regard to

---

**Figure 5** Spectrum of error noise for ETSI 34 Mbit/s codecs.

---

Dr Chris Dalton graduated from the University of Bristol (UK) in 1960 with a doctorate in Electrical Engineering and has been associated with Broadcasting since that time. He worked for 18 years with the BBC, in the Research and the Designs Departments, prior to joining Questech Ltd. as Technical Director. In 1982 he moved back into Broadcasting when he joined the Independent Television Association where, until recently, he was Head of Development at the ITV Network Technology Centre, in London. He is now retired but intends to maintain an active interest in the Broadcast industry.

Chris Dalton has specialized in colorimetry and the composite PAL system, and has been closely linked with a number of pioneering developments in digital studio systems. His current interests are new broadcast systems and standards and, in particular, extended- and high-definition TV and digital broadcast systems.

Dr Dalton participated in a number of international activities with the EBU and ITU. He is a member of the Institute of Electrical Engineers, a Fellow of the Royal Television Society, and a Fellow of the SMPTE. His achievements have been recognized with awards from both the IEE and the RTS.
the number of different types of codec that may be concatenated.

5. Conclusions

With the adoption of broadcast equipments that incorporate compression techniques into the production and transmission environments, it is important to consider the cumulative effects of concatenating compression codecs throughout the broadcast chain.

By considering alternative scenarios, it has been demonstrated that the piece-meal adoption of compression equipments of dissimilar types within existing large analogue or digital environments could lead to operational difficulties and the degradation of picture quality. The eventual adoption of an integrated scenario, using standardized MPEG-2 422P@ML codecs for studio applications and MPEG-2 MP@ML codecs for transmission, will result in the efficient application of compression-based products.

Objective methods of assessing a compression system are clearly required, ideally in conjunction with the tools to predict the overall performance of cascaded codecs within a broadcast network. Methods of making objective evaluations of equipments which incorporate compression have been discussed briefly, with an indication that techniques are currently in development which, in turn, will lead to the availability of instrumentation for objective measurements.

Bibliography


