RECOMMENDATION ITU-R BR.657-2*, **

Digital television tape recording

Standards for the international exchange of television programmes on magnetic tape***

(Question ITU-R 103/11)

(1986-1990-1992)

The ITU Radiocommunication Assembly,

recommends

that, for the international exchange of digitally recorded television programmes conforming to the 4:2:2 member of the family of standards (Recommendations ITU-R BT.601, ITU-R BT.656 and ITU-R BS.647), the technical criteria should conform to the parameters given in IEC Publication 1016****.

NOTE 1 – Operating practices for the D-1 format can be found in Recommendation ITU-R BR.779.

NOTE 2 – Annex 1 shows a basic functional block diagram for the D-1 format.

NOTE 3 – Annex 2 contains the basis for the digital television tape recording standard.

---

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

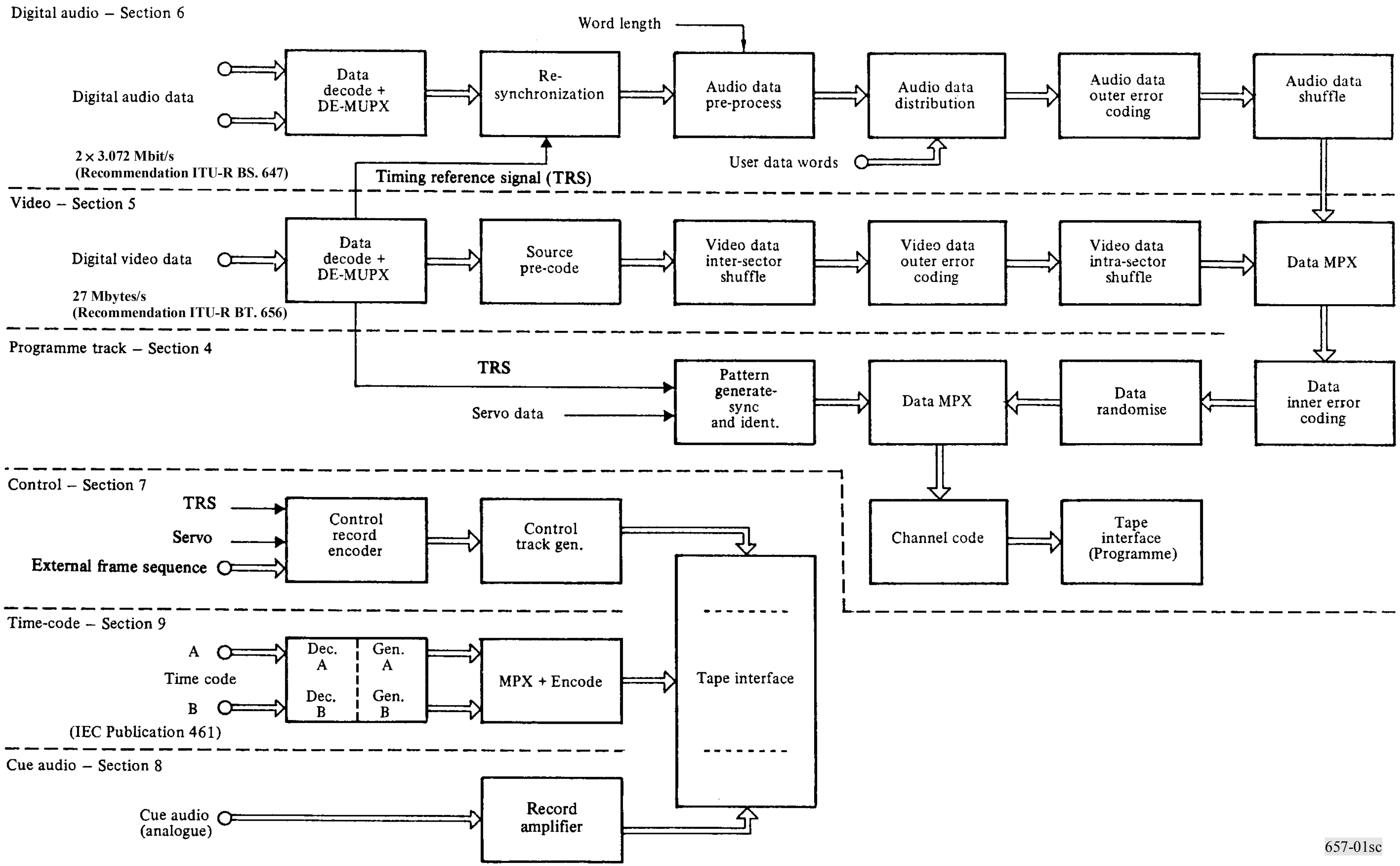
** This Recommendation should be brought to the attention of the International Electrotechnical Commission (IEC).

*** International programme exchange is defined as the transmission of television or sound programme material (or components thereof) among professional parties in different countries. It should be based on internationally agreed and widely employed technical standards or operating pratices, except by prior bilateral agreement among the parties involved.

**** A request for new work has been issued to incorporate a number of changes to the system specification.

ANNEXE 1

Functional block diagram 4:2:2 digital recorder record path processing



ANNEX 2

Basis for the digital television tape recording standard

Introduction

This Annex describes the background for the choice of parameters for the digital television tape-recording format specifications given in this Recommendation.

The specifications are based on inputs received by the ITU-R from a number of sources and notably from the EBU, the OIRT and the United States of America, the latter describing work carried out in the SMPTE.

The operating requirements on which the specifications are based were agreed among a majority of users in the bodies above, although some spread of opinions was observed among users.

The technological feasibility and viability of the specified format was confirmed in the same bodies, by means of consultations between the users and the manufacturers.

The text is structured in several sections as follows:

Section 1 – Users' requirements for digital television tape recorders

Section 2 – Parameters of the tape format

Section 3 – Mechanical characteristics of tape cassettes

Section 4 – Source coding parameters for the digital video and audio signals

Section 5 – Signal processing in the DTTR

Section 6 – Parameters of signals recorded on the longitudinal tracks

Section 7 – Explanation of terms

1 Users' requirements for digital television tape recorders

1.1 General requirements

1.1.1 A digital television tape recorder (DTTR) should record digital video signals according to the 4:2:2 standard specified in Recommendation ITU-R BT.601 and 4 digital sound signals according to the standard specified in Recommendation ITU-R BS.646 (sampling frequency of 48 kHz and at least 16 bits/sample linear coding). The timing relationship between the 4 digital audio channels should be adequate to permit any combination of 2 channels to be used for stereo pairs.

1.1.2 The DTTR should employ cassettes that will protect the tape from dust and other similar hazards. The cassettes should use reels with two full flanges. There should be a family of cassette sizes that can be used interchangeably in the full studio version DTTR.

1.1.3 The number of different cassette sizes should be kept to a minimum but should be adequate to cover the unique requirements of a series of recorders such as a production/post-production recorder, a portable recorder and a multi-cassette recorder.

The largest cassette size should be adequate to provide 76 min of record/play time with current 16 µm thick tape, so allowing for 94 min with 13 µm thick tape. The medium and small size cassettes should be able to contain 34 min and 11 min respectively, of 16 µm thick tape.

1.1.4 Consideration should be given in the design of DTTRs to the use of these recorders, and their initial application, in an analogue TV studio environment. Optional composite and/or component video analogue inputs and outputs shall be available to meet this need. Similarly, provisions shall be made for optional analogue audio input and output signals.

1.1.5 Two of the longitudinal tracks are user requirements. One track shall be capable of recording audio for use as a cue channel to aid in editing and the other is to be for recording a form of time and control code. They should be independent of the main digital audio and video channels and should be readable at shuttle and search speed over the range 0.1 to 50 times normal speed in both directions.

1.1.6 For the 525-line systems, lines 14 to 263 and 276 to 525 should be recorded and for the 625-line systems, lines 11 to 310 and 324 to 623.

1.1.7 Provision should be made for ancillary data to be included with the video and each of the audio signals; concealment should not be applied during field blanking when ancillary data may be carried.

1.1.8 It would be desirable if a DTTR built for a given level of the family of digital TV standards (see Recommendation ITU-R BT.601) were able to handle lower levels (at least, the replay of recordings made at lower levels).

1.2 Performance parameters in normal play mode

1.2.1 Assuming no uncorrectable errors due to the record/play-back process, the DTTR should be transparent with regard to the digital inputs as specified in § 1.1.1.

1.2.2 After 10 generations, essentially there should be no perceptible impairment to the audio and video signals, with critical programme material.

1.2.3 After 20 generations, the impairment should preferably not exceed 1/2 grade and certainly not exceed 1 grade of the 5-grade CCIR impairment scale. For further generations, the increase in impairment should be gradual.

1.2.4 The analogue editing audio channel should provide a bandwidth of the order of 10 kHz.

1.2.5 The accuracy of the recording and the replay of the digital audio and video information should be adequate to ensure that after 10 video and/or audio edits, or 10 video and/or audio generations, the accumulated relative timing error at any stage will be less than 40 ms.

1.2.6 The output time and control code signal on the longitudinal time and control track shall have a maximum timing error on interchange of ±1 ms with respect to the output video.

1.3 Operational requirements

The DTTR format shall offer the possibility of the same operational features and editing flexibility as the most advanced present-day recorders. In this respect, the most sophisticated production/post-production recorder shall be capable of at least the following features:

1.3.1 General features

– Broadcastable pictures at continuously variable speeds from about minus two times through still frame to about three times normal playing speed.

– Full quality picture and sound over a range of about 90% to 110% of normal playing speed. Audio should be full quality but not correct pitch unless an optional feature is fitted.

– Recognizable pictures at speeds from 0 to 20 times normal playing speed in both directions. The digital audio recovered in this mode shall have recognizable content and would have minimal change in pitch.

– Shuttle speeds from 20 to 50 times normal playing speed in both directions with major scene changes perceptible.

– Fully locked picture and sound in less than 1 s from standby (slack tape and head spinning) mode and instant start from still frame.

– It is a desirable option that video/audio simultaneous record/replay be available for confidence in record checking.

– It is desirable that the transport provide switchable 525/625 operation.

– Variable forward and reverse broadcastable fast motion up to a maximum of 6 times normal speed would be desirable.

1.3.2 Additional editing features

– Video editing with single field resolution and a minimum duration of one field.

– Insert and assemble modes.

– Independent editing of all channels (video, each of the four digital audios, analogue editing audio, longitudinal time code) and any combination of audio and video split edits in the same pass.

– Transfer of audio from any audio channel to any other with no delay introduced.

– An option is required that permits extracting the play-back audio digital bit streams from the DTTR in advance to compensate for external processing delays and re-recording it on an audio channel maintaining the original timing relationship.

– Video time code usable up to about 20 times normal playing speed in either direction.

– Operation by remote control using a standard machine control interface, such as the ES bus system developed by the SMPTE and EBU (EBU Document Tech. 3245 and Supplements).

– Digital audio editing with better than 6.7 ms resolution, with one field minimum insert duration and with overlap transitions of not less than 4 ms for the simplest recorder. For recorders with audio read-modify-write capability, the overlap duration should be adjustable to accommodate programme characteristics (4 to 50 ms).

1.4 Other requirements

1.4.1 The DTTR should be highly reliable and easy to operate.

Operating requirements are as follows:

– the DTTR should be designed to be operated by non-technical personnel with minimal training;

– the alignment controls required for routine operations should be minimum;

– the DTTR should operate reliably even in rather poorly controlled environmental conditions.

1.4.2 The DTTR should be easy to maintain. Maintenance requirements are as follows:

– the DTTR should be modular in concept to ease the identification of failed modules and to minimize the amount of realignment necessary after the replacement of a module;

– the DTTR should have indications to advise the operator (when possible) of out-of-limit conditions that could signify imminent failure; an example of this might be an abrupt increase in raw error rate;

– indicators should be provided to indicate a failure condition to advise the operator and maintenance personnel of the action to be taken, self-diagnostic or test routines should be provided to aid in isolating the failed module;

– the design of modules should be such that users can conveniently locate and replace failed components.

2 Parameters of the tape format

2.1 Basic assumptions

The track configuration is based on a number of assumptions and on the needs of users. The assumptions were that:

– the magnetic coating would be of the improved metal oxide type;

– on such a coating, the minimum recorded wavelength would be 0.9 μm;

– one wavelength would correspond to the recording of two bits;

– the number of lines recorded per television field would be 250 in the 525-line system and 300 in the 625-line system;

– the DTTRs would be of the helical-scan type;

– the total bit rate (corresponding to the video and audio signals together, recorded on the programme track with the appropriate protection, and the edit space between them) would be 277 Mbit/s;

– there would be a ratio of 5/6 between the number of tracks per field in 525 and 625-line systems (this assumption, implemented in conjunction with the previous assumptions, aims at allowing the common use, in 525 and 625-line DTTRs, of as high a number of elements as possible);

– the recording of one television field would be carried out on a total of 10 tracks in the 525-line system and 12 tracks in the 625-line system;

– the tape would be packaged in a cassette for programmes of a least one hour; the extension of the duration to one and a half hours would be foreseen.

Some of these assumptions were based on preliminary feasibility studies, and are briefly described below; others were arrived at, as optimal compromises, in the course of the definition of the recording standard.

2.2 Choice of helical scan recording

The high bit rate to be recorded on tape requires a very high writing speed; the data rate exceeds 200 Mbit/s, when unnecessary redundancy is removed, and necessary auxiliary and error protection signals are added. The application of some sort of multi-channel recording with stationary heads was considered inadequate, and therefore a system using rotary heads was the obvious choice. Previous experience of this kind of recorder has shown the major positive advantages of helical recording which was accordingly selected.

2.3 Choice of the magnetic material

Some theoretical studies, and practical experiments, have shown that metal particle tapes and more particularly metal evaporated tapes, can offer higher packing densities, than conventional oxide tapes. Extensive research and development in the field of "metal" tapes is under way, but as it seemed inadvisable to base the standardization on tape technology which has not been proven improved oxide tapes were the logical choice. It was found that a viable full quality professional digital video tape recorder is achievable with present technology, and that the advent of "metal" tapes in the future may offer an increase of the operational security margin.

2.4 Choice of the minimum recorded wavelength

At the beginning of the standardization process it seemed that 1 μm was the smallest practical value for the shortest wavelength to be recorded. It was also known that video heads could be manufactured for shorter wavelengths, and that these wavelengths offered better packing densities, although at shorter wavelengths the effects caused by drop-outs became more critical. Considerations of overall reliability led to the adoption of a value of 0.9 μm for the shortest wavelength.

2.5 Choice of the video tape width

One of the major issues was the tape width. Initially it was assumed that the optimum width would be 1 inch (25.4 mm), but it soon became evident that other dimensions were also practicable, and in some aspects perhaps even more suitable.

Eventually the discussion centred on the selection between 25.4 and 19 mm wide tapes. The final choice was based on the assessment of such critical parameters as:

– cassette playing time;

– guidability for the tape;

– forces involved at different points of the tape path;

– aspects of the portable DTTR;

– search time.

The longest cassettes were expected to offer 94 min playing time with 13 µm thick tape, and consequently 76 min with 16 µm thick tape. When a comparison was made between the dimensions of the cassettes for 25.4 mm and 19 mm tape width, it was found that differences in size, volume and weight were finely balanced and did not significantly favour any of the two proposed tape widths. However, the assessment of the behaviour of the two tapes on the machine transport indicated some important differences. The mechanical analysis showed that for a given thickness of tape, the guidability of the tape and the mechanical forces at some critical points on the tape path depend on tape width, and that a narrower tape offers advantages which become more significant when the tape thickness is reduced.

The penalties of adopting the narrower tape were judged of negligible significance to broadcasters but the advantages of the narrower tape, which could result in the same mechanism being usable in a range of recorders for a variety of applications, were considered to be significant.

These considerations led to the choice of 19 mm for the tape width.

2.6 Design of the track pattern

The track pattern was designed so that it would meet the following requirements:

– recording of the component digital video signal;

– recording of four independent digital programme audio signals;

– recording of a time and control code;

– recording of a control track;

– not to preclude the achievement of a broadcastable picture at speeds other than normal and a recognizable picture at shuttle speed;

– providing a "recognizable" audio at speeds other than normal;

– providing maximum 525/625 equipment commonality.

Three longitudinal tracks are provided, allocated to:

– the tracking control signal;

– the time and control code;

– an analogue "editing" or "cue" audio signal.

The transport for the DTTR employs a helical segmented format for video recording. For reasons of complexity and economics, the programme audio tracks are multiplexed with the video track but in such a way that video and all audio channels can be individually recovered and edited. The channel coding, data rate and format, and packing density, are identical for audio and video. The minimum recorder wavelength is approximately 0.9 µm in a 45 µm track pitch. There are 20 tracks per TV frame in 525 lines (24 in 625 lines) and audio bursts are recorded in duplicate. It was found convenient from error-ratio considerations to locate the audio data in the centre of the track. The audio data is written in two different positions in such a way that scratches, head failures and channel failure have a minimum effect. Gaps are provided to allow independent video and audio editing on tape, and it should be noted that each burst contains only audio from a single audio source. The arrangement lends itself also to a number of additional features in editing.

2.7 Editing

On-tape editing of video and audio has been stated to be an important feature of the DTTR by the users, who requested that each channel be individually editable, to the smallest possible increment. It should be noted that in addition to the editing capabilities of the DTTR itself, the digital recording process enables any data to be transferred to other editing systems (e.g. computer or disc based) processed and returned to the tape with minimum impairment, thereby allowing complex editing, improvements, etc., to be handled very effectively in conjunction with the DTTR.

The format proposed makes provision for several modes of operation.

2.7.1 Cut edits

At the edit point, the relevant sectors of the previously recorded programme are replaced by those of the incoming material, by gating in the record circuits during the appropriate time intervals. For video, this is the only envisaged mode and it provides a time increment of one field (but the off tape signals must be kept synchronous with the incoming video at frame rate). For an audio channel, an increment of four tracks (6.7 ms) is established, no processing is involved and the protection of the audio data is not affected. A transient may however be generated due to the very sharp transition between segments in play-back.

2.7.2 Simple overlap audio edits

At the beginning of the overlap period, the content of one of the two pairs of audio sectors is replaced by the new data without changing the other pair containing the old data. At the end of the overlap period both bursts are replaced. The new bursts written during the overlap period contain a flag to indicate the overlap. This edit method is very applicable to portable machines due to its basic record-mode simplicity, but audio is somewhat less secure during overlap due to the loss of redundancy. It has an increment of four tracks (6.7 ms).

2.7.3 Processed overlap audio edits

More elegant audio editing can be obtained by performing a read-modify-write operation on the audio sectors, using an advanced-read head to ensure that the modified data bursts are returned to the tape at the correct locations. Due to the digital nature of recording, no impairment is introduced by this operation. The resolution of this method of operation is theoretically one sample or 20 µs. The additional complexity to perform edits of this nature will likely limit its application to studio level machines.

3 Mechanical characteristics of tape cassettes

3.1 The users' requirements

Outlining their views on the future digital video tape recorder, the users stated that an open reel machine might be acceptable for the "first generation" of digital machines, but that the ultimate goal should be a cassette configuration. The need to protect the tape as much as possible from ambient dust and handling stresses (which could considerably increase the drop-out activity) made the cassette principle the only possible approach for a general purpose digital television tape recorder.

The users also pointed out their expectation that the future digital recorder should become available not only as a studio (or OB) machine, but also as a multi-transport machine for short programme segments, and, in a more distant future, as a portable recorder. In order to provide for all these needs three cassette sizes were selected, and completely mechanically defined:

– small (S),

– medium (M),

– large (L).

3.2 The design of the cassette

The starting point of the new family of cassettes was the existing 8 mm cassette design. It was decided that for professional use reels with two flanges were mandatory in the cassette.

The design of a new tape cassette for professional use offered the possibility of implementing some specific features like programmable "holes". Four holes in the base plate of the cassette would be at the manufacturers' disposal and used to indicate features like tape coating material, thickness, etc. Four additional holes on the same plate would be reserved for users, for "record inhibit" and similar functions. The position of the holes should allow their detection when cassettes of different size (S, M and L) are played on the same machine.

Since it is considered that standardization of the mechanical characteristics of cassettes is a task for the IEC rather than the ITU-R, the present ITU-R Recommendation on digital television tape recording does not go into the details of cassette standardization, but refers the reader to available documentation, pending a formal standard to be issued by the IEC.

4. Source coding parameters for the digital video and audio signals

4.1 Source coding of the digital video signals

The starting point for the complete standardization process is the requirement that the DTTR should be able to accept at its input and deliver at its output digital component video signals in full conformity with Recommendation ITU-R BT.601. The interface is in accordance with Recommendation ITU-R BT.656.

The DTTR records only 300 lines (625/50) or 250 lines (526/60) per field. Most of these lines carry picture information, but the remainder may carry ancillary data information and in the play-back mode they should not be subject to error concealment techniques which should only be applied to the active picture area. Only 1 440 samples of the active line are recorded.

4.2 Source coding of the digital audio signals

The audio input and output signals complying with Recommendation ITU-R BS.647 are serial data streams which may carry two audio signals (e.g. a stereo pair), each with its own status and user data.

A minimum of two such data streams is required to feed the four channels of the DTTR. However, there may be applications where individual data-streams-per-channel are required, with the second signal in each unused.

The capacity of each data stream corresponds to two 24 bits/48 kHz audio signals, each with a 48 kbit/s status channel and user and housekeeping channel containing, e.g. sample validity, parity and synchronization bits. There may also be some applications where analogue signals are directly encoded at the DTTR and in this case only the audio data will be present.

Four audio connectors, numbered 1 through 4, are provided for the application of individual sound programmes to the four channels of the DTTR. However, connectors 1 and 3 may also be used for stereo pairs.

In the case of monophonic programme sound, this should be carried on digital audio channel number 1.

In the case of stereophonic programme sound, the left and right channels should be carried on digital audio channels numbers 1 and 2, respectively. This stereo pair may be applied through connector 1.

If additional programme sound components are needed, they should be recorded on digital audio channels numbers 3 and 4. If these components are a stereo pair, they may be applied through connector 3.

It is possible to satisfy almost all possible applications and practices, and still preserve the necessary compatibility by selecting eight different modes of organizing the 20-bit audio words obtained by rounding off the original 24-bit words.

In these eight modes the length of the audio word varies from 16 bits (with one status, one user, one validity and one unassigned bit) to 20 bits when only audio data is present (in the case, for example, when the analogue audio is directly encoded at the recorder input). In the play-back mode the audio data are re-formatted into the format of Recommendation ITU-R BS.647 so that the output is normally identical to the input signal.

5 Signal processing in the DTTR

5.1 Outline of the record and play-back processing

Digital audio data is multiplexed in blocks with video data to obtain a high packing density and to take advantage of the economies of common error correction, heads, read/write amplifiers, clock recovery, etc.

Annex 1 to this Recommendation shows a conceptual block diagram of the digital treatment of video and audio.

The mechanism of saturation recording on magnetic tape is essentially simple but the signal processing required to use this recording mode in the most efficient way is relatively complex due to the need for effective control of the resulting data errors at the packing density required. On the record side of the DTTR, the processor must assemble blocks of words representing video, audio, status/user data and internal control data, and add to them the necessary redundant words to allow very secure detection of word errors and a good level of error correction, invoking error concealment when correction overflows. The processor must also add the necessary synchronizing information and block identification to allow block recovery and orderly reassembly of the data stream. The data is coded into a recording format which has appropriate spectral characteristics for the actual channel used, and also includes a strong clock recovery capability. In this process, the sequence of video or audio words is shuffled, so that adjacent samples of the input signals are separated and well spaced on the tape. This permits more effective concealment when burst errors occur. Finally, the record processor outputs the data in burst mode to different heads so duplicating the audio blocks on two separate tracks. This additional spatial redundancy greatly improves the probability of successful recovery of the data in the presence of major errors caused by tape scratches or head-clogging and also provides some useful edit features. By the time it is recorded on tape the data has grown by about 290% compared to the original data at the input to the recorder.

To simplify the design of the recorder, part of the error-correction processor and most of the sync and clock processing, channel coding and read/write logic for the audio channels can be integrated with that of the video channel.

The recovery of the data from the tape is the inverse of the record processor-channel, i.e. decoding, sync recovery, identity check, error detection, correction and concealment, and demultiplexing into the various streams for the output processor and the internal DTTR controls. While audio or video data can be concealed (interpolated) if uncorrectable errors are detected, such is not the case for status or user data, or for control words and these must be processed differently. The output processor retimes the data and reassembles the original data stream of video samples, audio samples, status, user information and sync data, and fills the null areas where no data is available such as in the four LSBs of the audio word, dropped in the input round-off. Except for these bits the output signals are a precise copy of the input except during infrequent concealments, consequently numerous generations can be made without the accumulation of impairments.

5.2 Error control

Data recovered from the tape is impaired by a number of artifacts added during the record and play-back process:

– random errors due to noise, interference, tracking imperfections;

– burst errors due to head/tape contact failures, tape drop-outs and tape roughness;

– large burst errors due to failures such as tape scratches, head clogging, channel failures.

As the objectives established for the DTTR include an audio quality grade of 4.5 on the ITU-R 5-point scale after about 20 generations (i.e. after about 20 generations, one-half of a group will be unable to hear any difference compared to the original) errors must be eliminated to a very large degree, and in such a way that the burden placed on the DTTR channels is minimized. A further complication is that the most economical arrangement of the DTTR is achieved if there is a maximum of commonality between the video and audio channel hardware, bearing in mind that the audio represents only 2% of the total data, but requires a final error rate about 100 times better than video. In addition, both the video and audio data are autocorrelated (i.e. there exists an implied relationship between adjacent samples), and so missing or damaged samples can be replaced by an approximation derived from adjacent samples, while the status, user and control data must be considered random and hence cannot be estimated in the general case. This may result in different error objectives for audio, video, and data in the same data stream. Clearly, error control is a very important factor in the design of the DTTR audio system.

Based on the above considerations and taking into account that:

– the code must provide near-perfect detection of errors;

– the code must add a minimum of overhead;

– the expected error statistics are known;

– commonality of coding of the audio channels with the video channel is desirable;

the choice was made of a Reed-Solomon product code based on a common (60 + 4) bytes inner code in Galois Field 256 (GF(256)). The inner code provides the basic protection against short duration random error sources, such as noise or short drop-outs, and enables it to correct these errors. However, the same code should also serve to reliably detect more extensive error sources such as long drop-outs and scratches since these can be suitably processed by the outer code.

The inner code also requires to be active during replaying at shuttle speed. The number of errors is very high in such circumstances and is likely to overload any reasonably complex correction code. Allowance must therefore be made for the use of concealment.

For video, the outer code block size is set at 30 data bytes plus 2 Reed-Solomon check bytes in GF(256) to give a product block which is (60 + 4) by (30 + 2). Ten such product blocks yield the total array having a row dimension of (600 + 40) bytes and the column dimension of 30 bytes with 2 check bytes. During the recording, the inner code blocks are sequentially written on tape, one row at a time. In play-back, the inner code blocks are normally decoded first.

The data corresponding to successive picture elements in the television line, which arrive at the recording heads after being spread into blocks and accompanied by protection data, are recorded on four successive sectors in order to facilitate the protection strategy by distributing the effects of head failure.

In order to deal with error bursts corresponding to extended drops in level, the product code uses the inner code to determine the locations of the drop-out by employing the error detection capability of the inner code. Once the location of the drop-out is found, then the outer (or vertical) code is used to correct the drop-out error. This outer code is, in effect, through the action of the product code, operating on words which have been interleaved to a depth of 600 bytes.

Since the outer code can correct any two rows known to be in error, the maximum correctable drop-out length is 1 200 bytes (equivalent to 4.8 mm of track length). Further, the outer code provides for double error correction, and consequently the correction of multiple short bursts, guaranteeing the correction of all double drop-outs up to 600 bytes in length. Multiple bursts beyond two in each product block can be corrected but correction is not guaranteed as it depends on the drop-out lengths and locations.

In order to reduce the effect of uncorrectable drop-outs and scratches which generally run along the length of the tape, and to improve pictures in shuttle, the distribution of video data words in each of the four recording channels is completed by a shuffling along each video sector.

Without shuffling, a scratch or roughness resulting in a large drop-out would be likely to cause, in a part of a picture segment, a simultaneous local loss of information from two of the four heads. In the case of a scratch this would repeat in every picture segment and from field to field. Since an uncorrected error tends to be very much more visible than a concealed error, when the error correction is overloaded the best approach is to conceal all the words that are reasonably suspect.

Concealment can be best achieved when any errored word is well isolated from other words in error. However, the better the isolation, the lower the number of errors than can be concealed. It is necessary, therefore, to arrange, as far as possible, that as the word error ratio increases, the errors are spread uniformly and do not cluster in parts of the picture, since this would make error concealment impossible.

The algorithm chosen for the shuffling has the characteristic that as the drop-out length increases so does the density of errors, but the density will always be substantially uniform throughout the affected segment of 50 lines.

Under normal play conditions, concealment will be used relatively infrequently but during shuttle the situation is totally different and the words requiring concealment may exceed the number of correct words. If the loss of information were substantially equal in all segments, the resulting shuttle picture would be more than adequate for editing purposes. However, at certain critical shuttling speeds, the loss of information could vary significantly between the segments and the loss of information could repeat from field-to-field if the same shuffling were used. The four-field variation of the shuffling sequence, provided by the algorithm, decreases the incidence of critical shuffling speeds.

For audio, the product code is based on a (60 + 4) inner code, common with the video channel, and a (7 + 3) Reed-Solomon GF(16) outer code. This provides the necessary burst-error correction. This coding is backed up with full duplicate writing on the tape, to overcome major faults and to give powerful correction of burst errors. Given the error statistics of the channel, a concealment rate of one or two per minute is anticipated for the audio at the 20th generation, providing very acceptable levels of performance. Undetected errors are at a negligible ratio. The audio data is shuffled in the block prior to writing on tape to improve error concealment over 6.7 ms. Based on these error-correction methods, the DTTR is expected to provide audio performance limited only by the selected word length and the performance of the initial A/D coder and filter, for many generations, and thus providing a high level of technical transparency.

5.3 Tape data format

After the useful data has passed through the outer error coding, shuffling, interleaving and inner error coding, it is arranged in blocks of fixed length, corresponding to one row of the inner coding. By the addition of sync and identification (ID) information, it is converted to a sync block, the smallest unit of data recoverable from tape. This is then passed through the channel coder to prepare it for the head-to-tape interface. Sync words are of identical structure for video and audio blocks. 160 sync blocks are included in a video sector and 5 sync blocks are in an audio sector. Sectors start with a preamble sequence and end with a postamble sequence. Sectors are separated by an unrecorded edit gap to allow some positional tolerance. Audio sectors are written on to the tape at two locations using different heads to improve the probability of successful recovery.

The channel coder, common to all data written by the rotating heads, modulates the channel with the data-stream in a manner that improves data reliability by spectral shaping (e.g. elimination of d.c. and low frequency components) and eases clock recovery in play-back over the speed range of interest.

Data recovery is a complementary process to that previously described, i.e. channel decoding, clock and data recovery, sync and ident. recovery followed by inner error detection and correction. Up to this point the video and audio share the same path. Subsequent processing is performed separately, i.e. deshuffling, outer correction followed by concealment of any residual errors that are detected but not corrected.

6 Parameters of signals recorded on the longitudinal tracks

6.1 Cue audio track

In editing operations there is a need for audio recovery to be intelligible over a wide range of speeds and it is clear that the digital tracks using burst techniques cannot provide this capability in a simple manner. A longitudinal editing track is thus included in the format and for the sake of simplicity conventional a.c.-biased analogue recording is specified with a track width of about 600 μm. Analogue recording does not overcome the distortion and print-through problems due to the very thin coating and base thickness (13-16 μm) used for the digital recording media, but performance at variable speed is better for a given complexity level and is adequate for the purpose of providing approximate points for editing.

6.2 Time-code track

For reasons similar to those described for the longitudinal cue audio track, a time-code track is included to carry video-related time-code for edit control and scene access.

It is noteworthy that the four digital audio channels each carry a double time-code in their status bits and hence a total of nine time-codes and user bits may be present in the DTTR.

Studies to record additional time-code information within the existing data capacity of the format, such as user bits of the time, are in process.

6.3 Control track

The control track modulation is 3-state and consists of pulse doublets separated by mid-level intervals and has an average d.c. component of zero.

The servo reference doublets occur every two video segments, that is 5 times per frame for 525-line systems, 6 times for 625-line systems; they have a nominal frequency of occurrence of 150 Hz. An additional doublet occurs once per television frame to provide a frame reference.

Since there will be 1 601.6 audio samples per 525-line frame, giving 8 008 samples per 5 television frames, an additional doublet is used to mark the control track every 5 television frames. For 625 lines there are 1 920 audio samples per frame, so this pulse doublet is not required.

An additional pulse doublet has been defined to provide a reference for editing video frames in proper sequence. In addition, this pulse doublet can be used to indicate colour frame start, should this be required, for synchronizing the DTTR to an external colour reference.

The period after the end of this optional doublet, up to the start of the next servo reference doublet, is the time when an edit may occur and is reserved for this purpose.

6.4 Timing relationships

In a practical analogue machine, the timing relationships at the input and output must be specified, and usually the audio and video are time coincident. The timing relationships on the tape are specified to take account of the physical constraints of head placement and to minimize the need for compensating delays, particularly on the record size. In the case of the digital recorder, further complications exist due to the timing relationship between the audio and video sampling clocks, the use of burst-mode operation for the audio, multiplexed into the video channel and the use of interleaving and shuffling to improve error correction and concealment.

The DTTR will follow conventional practices and have the audio and video coincident at the input and output with time-coincident bursts of audio and video data in the same tracks. Cue audio and time-code on longitudinal tracks are offset 210 mm from the corresponding digital tracks.

7 Explanation of terms

7.1 General definitions

7.1.1 *Programme area*. The programme area is that part of the tape on which is recorded the programme digital video and digital audio signals.

7.1.2 *Programme area track pattern – Video and audio sectors*. A head which is recording during an entire scan of the programme area lays down a helical track consisting of six sectors of digital video and digital audio in the sequence video-audio-audio-audio-audio-video. 20 such tracks in the 525 system and 24 in the 625 system contain a video recording equivalent to the period of two television fields and audio recordings corresponding to 33.37 ms in the 525-line and 40 ms in the 625-line system for each of the audio channels. The recordings of a television field, however, commence at the start of a video segment.

7.2 Track pattern allocation – Video and audio segments

7.2.1 *Video segment*. A video segment contains the digital video data originating from one fifth (in the 525-line system) or one sixth (in the 625-line system) of a television field, and comprises four video sectors. These are located in four adjacent helical tracks being the upper adjacent video sectors in the first pair of tracks and the lower adjacent video sectors in the second pair of tracks.

7.2.2 *Audio segment*. An audio segment initially contains the digital audio originating from a 6.7 ms period of an audio channel and comprises four audio sectors, distributed among four adjacent tracks. Hence, the four audio segments corresponding to a given time period are associated with two video segments corresponding to the same time period, and are physically recorded at the end of the video segments.

7.3 Electrical signal allocation

7.3.1 *Video and audio sector allocation – preamble, sync block, postamble*. Each video sector consists of a preamble, 160 sync blocks and a postamble. Each audio sector consists of a preamble, five sync blocks and a postamble.

7.3.1.1 *Preamble*. A preamble consists of a run-up sequence, a sync pattern, an identification pattern and a fill sequence.

7.3.1.1.1 *Run-up sequence*. A run-up sequence consists of a sequential bit pattern chosen to facilitate the locking of data extraction circuits.

7.3.1.1.2 *Sync pattern*. A sync pattern consists of two consecutive bytes whose bit pattern is chosen to be a robust indication of the start of a sync block.

7.3.1.1.3 *Identification pattern*. An identification pattern consists of four consecutive bytes, providing a unique address of the position of a sync block within four fields of recorded data, coded such as to remove d.c. and provide error protection.

7.3.1.1.4 *Fill sequence*. A sequence of bytes whose purpose is to maintain clock synchronization and not to carry useful data.

7.3.1.2 *Sync block*. A sync block consists of a sync pattern followed by an identification pattern followed by two inner code blocks.

7.3.1.3 *Inner code block*. An inner code block consists of 60 bytes of video data, audio data or outer code check data, followed by four bytes of inner code check data.

7.3.1.4 *Postamble*. A postamble consists of a sync pattern followed by an identification pattern.

7.4 Sub-sets of binary data

Usually, for convenience in parallel digital processing, binary information is processed in groups of bits referred to in the literature as words and bytes. These terms have generally understood meanings but are not unambiguously defined. For the purpose of this terminology the following definitions are assumed.

7.4.1 *Byte*. A byte consists of 8 bits of binary information. It may have an identity other than being a convenient processing unit (see for example video data word), but generally this is not implicit.

7.4.2 *Video data word*. A video data word is a byte in which the 8 bits represent the possible 256 quantum levels of a video sample.

7.4.3 *Audio data word*. An audio data word consists of 20 bits. In the most basic operating mode, 16 bits represent the possible 2^{16} quantum levels of an audio sample and 4 bits are used for auxiliary signals. Other modes are defined in which one, two, three or four of the auxiliary bits are used to extend the dynamic range of audio sample quantization.

7.5 Error protection strategy

Various methods are used to reduce the effect of digital errors on the objective and subjective quality of the replayed video or audio.

The appropriate combination of methods to achieve an optimum result is generally known as the error protection strategy.

7.5.1 *Error correction*. The use of mathematically related check data recorded with the video and audio data, to locate and correct digital errors.

7.5.2 *Error concealment*. The replacement of erroneous samples by estimate values derived from related error-free samples.

7.5.3 *Source pre-coding*. The transcoding of video data words, so that for the most probable distribution of digital errors, there is a reduction in the peak error produced in a video sample.

7.6 Error protection – data organization

Error correction for both video data and audio data is of the product block type in which each data word is included in the computation of two sets of check data known as outer code check data and inner code check data respectively.

Additionally the video and audio data are redistributed from their naturally occurring sequences in order to reduce the effect of burst errors.

7.6.1 *Video data sector array*. For the application of product block error correction, the 18 000 video data words to be recorded in a video sector are considered as a rectangular array with a row dimension of 600 video data words and a column dimension of 30 video data words.

7.6.1.1 *Video outer code check data – video outer code block*. Video outer code check data consist of two bytes computed from a column of the video data array and regarded as being appended to that column. The resulting 32 bytes are known as a video outer code block.

7.6.1.2 *Video inner code check data – video inner code block*. Video inner code check data consist of four bytes computed from a 60-byte sub-set of a row of the video array (or a row of the video outer code check data) and appended to that sub-set. The resulting 64 bytes are known as a video inner code block.

7.6.1.3 *Video product block*. The array defined by 32 video inner code blocks and the corresponding 60 video outer code blocks is known as a video product block. There are 10 such video product blocks in a video sector.

7.6.2 *Audio data array*. An audio sector contains either odd audio data words or even audio data words. For the application of product block error correction, the 168 words of 20 bits each to be recorded in an audio sector are considered as a rectangular array with a row dimension of 120 words of four bits and a column dimension of seven 4 bit words.

7.6.2.1 *Audio outer code check data – audio outer code block*. Audio outer code check data consist of three 4-bit words computed from a seven 4-bit word column of the audio data array and regarded as being appended to that column. (In practice the audio outer code check data are distributed within the column.) The resulting ten 4-bit words are known as an audio outer code block.

7.6.2.2 *Audio inner code check data – audio inner code block*. Audio outer code check data consist of four bytes computed from a row of the audio array (or the appended audio outer code check data). The resulting 64 bytes are known as an audio inner code block.

7.6.2.3 *Audio product block*. The array defined by the 10 audio inner code blocks or by the corresponding 60 audio outer code blocks, is known as an audio product block. There is one audio product block in an audio sector.

7.6.3 Data redistribution for video and audio

7.6.3.1 *Interleaving*. The systematic re-ordering of data so that originally adjacent words of video or audio are separated, thus reducing the effect of burst errors on the error-correcting capability. The separation in words is known as the interleave distance.

7.6.3.2 *Shuffling*. The systematic re-ordering of video or audio data words to increase the probability that uncorrectable words are surrounded by error-free data words, for the application of error concealment.

7.7 Other electrical definitions

7.7.1 *Channel coding*. The process by which binary information obtained from the digital logic circuits, used in the processing of video and audio data, is converted to a waveform suitable for the recording on to a magnetic medium.

7.7.2 *Randomization*. The reduction of correlation in a serial bit sequence so that it statistically approximates to a random sequence.

7.7.3 *Scrambling*. Alternative term of randomization.

7.7.4 *Transcoding*. The recoding of data, by computation look-up table, so that there is a defined one-to-one relationship between each original code word and the derived code word.

7.8 Mechanical terms

7.8.1 *Basic dimensions*. A basic dimension is a fundamental dimension to which no tolerance is applicable.

7.8.2 *Derived dimension*. A derived dimension is obtained from other fundamental dimensions by computation and is given for reference purposes only.

7.9 Definitions related to editing

7.9.1 *Edit gap*. The space between adjacent sectors, to which edit transitions must be confined, between the end of the trailing sector postamble and the leading sector preamble.

7.9.2 *Cue audio track*. The longitudinal track reserved for the recording of analogue audio frequency signals which are to be used for production purposes.

7.9.3 *Control track*. The longitudinal track consisting of up to four sets of pulse doublets. Used for servo reference, indication of video frame, start of five-frame audio sequence (in 525/60 system) and may indicate, when required, the start of a colour frame sequence.