INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.188

(07/2002)

SERIES J: CABLE NETWORKS AND TRANSMISSION OF TELEVISION, SOUND PROGRAMME AND OTHER MULTIMEDIA SIGNALS

Miscellaneous

A framework for an efficient parallel video transmission system including codecs with functions of failure detection and picture quality evaluation

ITU-T Recommendation J.188

ITU-T J-SERIES RECOMMENDATIONS

CABLE NETWORKS AND TRANSMISSION OF TELEVISION, SOUND PROGRAMME AND OTHER MULTIMEDIA SIGNALS

General Recommendations	J.1–J.9
General specifications for analogue sound-programme transmission	J.10–J.19
Performance characteristics of analogue sound-programme circuits	J.20–J.29
Equipment and lines used for analogue sound-programme circuits	J.30–J.39
Digital encoders for analogue sound-programme signals	J.40–J.49
Digital transmission of sound-programme signals	J.50–J.59
Circuits for analogue television transmission	J.60–J.69
Analogue television transmission over metallic lines and interconnection with radio-relay links	J.70–J.79
Digital transmission of television signals	J.80–J.89
Ancillary digital services for television transmission	J.90–J.99
Operational requirements and methods for television transmission	J.100–J.109
Interactive systems for digital television distribution	J.110–J.129
Transport of MPEG-2 signals on packetised networks	J.130–J.139
Measurement of the quality of service	J.140–J.149
Digital television distribution through local subscriber networks	J.150–J.159
IPCablecom	J.160–J.179
Miscellaneous	**J.180–J.199**
Application for Interactive Digital Television	J.200–J.209

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation J.188

A framework for an efficient parallel video transmission system including codecs with functions of failure detection and picture quality evaluation

Summary

Parallel transmission links including codecs are often set up to realize a highly reliable contribution and primary distribution of television programs, in particular of big events such as Olympics and World Cup Soccer. This configuration, however, has the following issues:

1) The reserve link is totally wasted during normal periods.

2) Switching to the reserve link should be perfect to avoid a disrupted television program, when a failure occurs in a transmission link.

3) In addition, objective assessment of received video quality has been strongly demanded in an actual operation of digital television program transmission.

This Recommendation, therefore, recommends a framework for methods that support solutions for these issues. This framework mainly consists of the following three elements:

1) It reduces coding noise and improves picture quality compared to only one link by outputting the average signal of two links, when both links are normal (Efficiency).

2) It detects a failure that occurs in one of the two links by monitoring the two links by comparing them (Support of failure detection).

3) It evaluates the quality of the two decoded pictures by comparing them (Automatic picture quality assessment).

Appendices show examples of methods in this framework. Such methods improve the operational efficiency and reliability of parallel video transmission links.

Source

ITU-T Recommendation J.188 was prepared by ITU-T Study Group 9 (2001-2004) and approved under the WTSA Resolution 1 procedure on 29 July 2002.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

ITU-T Recommendation J.188

A framework for an efficient parallel video transmission system including codecs with functions of failure detection and picture quality evaluation

1 Scope

This Recommendation considers the parallel transmission link that is often used for contribution and primary distribution of television. Particularly, it considers the case where the parallel link includes compression encoders and decoders such as MPEG-2. Signal-to-noise ratio is improved by reducing the compression coding noise through averaging of the two links, i.e. the utilization efficiency of the links is improved. In addition, automatic detection of transmission failure is supported through monitoring based on a comparison of videos over the two links. Further, the assessment of picture quality such as estimation of signal-to-noise ratio is made possible by utilizing the difference of coding noise in two links. This Recommendation recommends that framework. The appendices describe examples of specific methods. Such systems are for use by organizations such as communication companies or cable operators.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

2.1 Normative reference

– ITU-T Recommendation J.143 (2000), *User requirements for objective perceptual video quality measurements in digital cable television*.

2.2 Informative references

– ITU-T Recommendation J.144 (2001), *Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference*.

– ITU-T Recommendation H.262 (2000), *Information technology – Generic coding of moving pictures and associated audio information: Video*.

– ITU-R Recommendation BT.800-2 (1995), *User requirements for the transmission through contribution and primary distribution networks of digital television signals defined according to the 4:2:2 standard of Recommendation ITU-R BT.601 (Part A)*.

3 Terms, definitions and acronyms

This Recommendation does not use any unusual terms or abbreviations.

4 User requirements

The user requirements for parallel transmission links that the framework described in this Recommendation should fulfil are the following.

4.1 Improvement in efficiency of transmission link usage

The reserve link is totally wasted in parallel transmission links during normal periods that occupy most of the time. The efficiency of transmission link usage should be improved by somehow utilizing also the reserve link during normal periods.

4.2 Improvement in reliability by real-time failure detection

It is important, when failure occurs, to switch to the reserve link as soon as possible to reduce the time during which disrupted pictures are output. Usually, operators monitor the link and switch to the reserve link when failure occurs. It is required to reduce the burden on operators and the time during which disrupted video is output, by automating this monitoring.

4.3 Improvement in efficiency of operation and maintenance by objective picture quality assessment by use of only received pictures

The degradation due to compression coding changes according to picture content even when there is no transmission error. The quality of received pictures should be automatically monitored. See ITU-T Rec. J.143.

5 Recommended framework

This Recommendation recommends a scheme for efficient parallel transmission of television.

5.1 General configuration

Figure 1 shows the framework recommended in this Recommendation to fulfil the requirements.

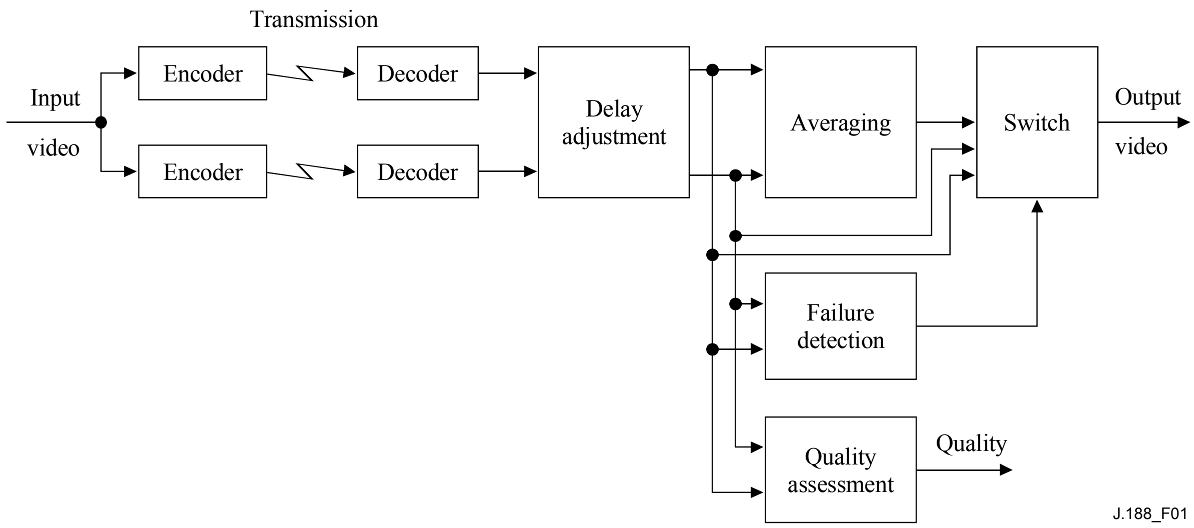

Figure 1/J.188 – Recommended framework

This Recommendation assumes that codecs are included in each of the two links.

First, one of the two received videos is, if necessary, delayed appropriately so that the transmission delays of both links would be the same. After that, averaging of two decoded pictures is performed to reduce noise, which leads to efficient usage of links, while failures are detected by comparing the two decoded pictures. Appropriate switching is performed according to this failure detection information. In addition, when there is no transmission error, objective quality assessment of the decoded pictures is performed by comparing the decoded pictures in the two links.

5.2 Description of each element

5.2.1 Adjustment of delay difference

This framework requires that the delay of two videos should be the same for the following averaging and quality assessment processing. It is recommended, therefore, that a delay adjustment unit should be inserted (Figure 1). Specific methods may be reading time codes or matching of pictures.

5.2.2 Averaging

The improvement in the efficiency of parallel link usage assumes that each of the links includes a compression encoder and a decoder (Figure 2). It is recommended that the two decoded video signals be averaged to cancel coding noise so that an improvement in signal-to-noise ratio could be realized. For this, it is important to reduce the correlation between the two noise signals due to the compression coding. Such methods include shifting video signals spatially and/or temporally, which is described specifically in Appendix I.

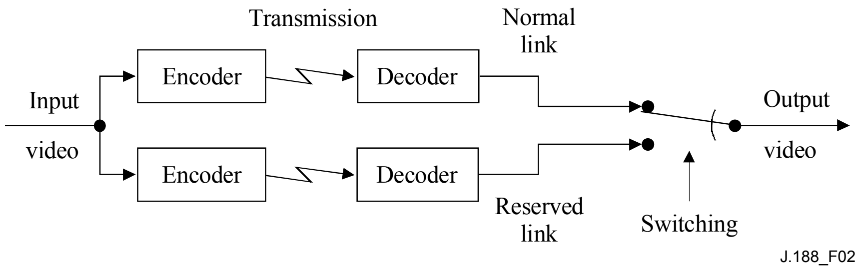

Figure 2/J.188 – Method for improvement in transmission efficiency

5.2.3 Failure detection

The parallel links considered in this Recommendation provide two received pictures. The recommended framework takes advantage of this characteristic to perform precise failure detection comparing to the case where only one received signal is available, i.e. the occurrence of failures is detected by comparison of two received pictures. Appendix II describes the specific methods.

5.2.4 Objective picture quality assessment

Picture quality assessment, as in 5.2.2, assumes a compression encoder and a decoder in each of the two links in a parallel system. It is recommended that the degradation of picture quality due to compression coding be estimated by utilizing the difference of two coding noise signals. Appendix III describes the theoretical background and a specific estimation method.

Appendix I

Averaging

I.1 Theoretical background

For improvement of efficiency of parallel links usage, we can reduce the correlation between the coding noise signals of two decoded videos to increase the improvement of the SN ratio by averaging.

Let x be the original signal, x_1, x_2 the two decoded signals, and d_1, d_2 their coding noise, respectively, where $x_1 = x + d_1$, $x_2 = x + d_2$ (Figure I.1). The noise power is supposed to be the same in the two channels and is represented by n. ρ_{d1d2} represents the correlation coefficient of d_1 and d_2.

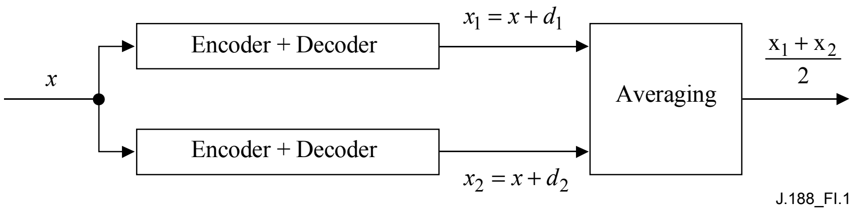

Figure I.1/J.188 – Averaging of two received signals

The coding noise d_m of the average signal of x_1, x_2 is:

$$d_m = \frac{x_1 + x_2}{2} - x = \frac{d_1 + d_2}{2}$$

and its power n_m is:

$$n_m = \sigma^2\left(\frac{d_1 + d_2}{2}\right) = \frac{n(1+\rho_{d1d2})}{2}$$

When the noises of the two channels are independent, therefore, the averaging improves SNR by 3 dB. Conversely, as their correlation becomes larger, the improvement becomes smaller. From experimental results, it has been found that an improvement of PSNR by 1 to 2 dB is realized.

I.2 Picture shifting method

In compression codecs, pictures are usually separated into many blocks and compression is performed block by block by utilizing spatial and temporal correlation. In order to reduce the correlation between the coding noise of two links, therefore, it can be considered that the picture is relatively shifted in at least one of the links. The shifting methods include a spatial method and a temporal method.

I.2.1 Spatial shifting method

Major picture coding schemes such as MPEG-2 (ITU-T Rec. H.262) perform an orthogonal transform block by block with the typical size of 8 × 8, and quantize the coefficients. By shifting the block boundary in one of the links, therefore, we can make the coding noise of two links different. Specifically, we shift the picture horizontally/vertically by a certain number of pixels. This picture shifting can be represented as follows. The original picture is represented as a one-dimensional signal for simplicity. *N* denotes the number of horizontal samples.

$$x(0),x(1),\ldots,x(i),\ldots,x(N\text{-}2),x(N\text{-}1)$$

The signal after horizontal shifting of *n* pixels is represented as follows.

$$x(N\text{-}n),\ldots,x(N\text{-}1),x(0),\ldots,x(N\text{-}n\text{-}1)$$

In other words, the pixels that are shifted out of the right edge are inserted into the left edge.

Figure I.2 shows an example where the picture is shifted horizontally by four pixels. In this figure, the part Q, which is shifted out of the right edge due to the right shifting, is inserted into the left edge. Figure I.3 shows the configuration of this spatial-shifting method.

1) In the processing prior to transmission, when the video is distributed to two links, the picture in one link is shifted horizontally or vertically by a certain number of pixels.

2) In the processing after transmission, the original position is recovered by shifting the picture in the inverse direction. Then, the pixel values of the videos in both links are averaged.

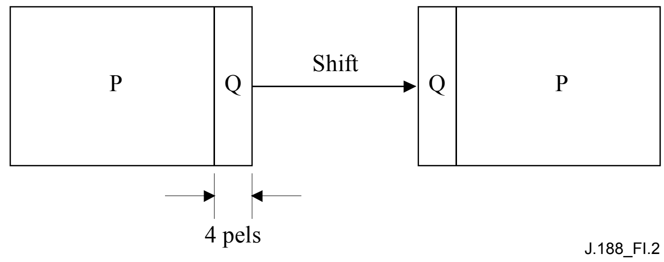

Figure I.2/J.188 – An example of spatial shifting

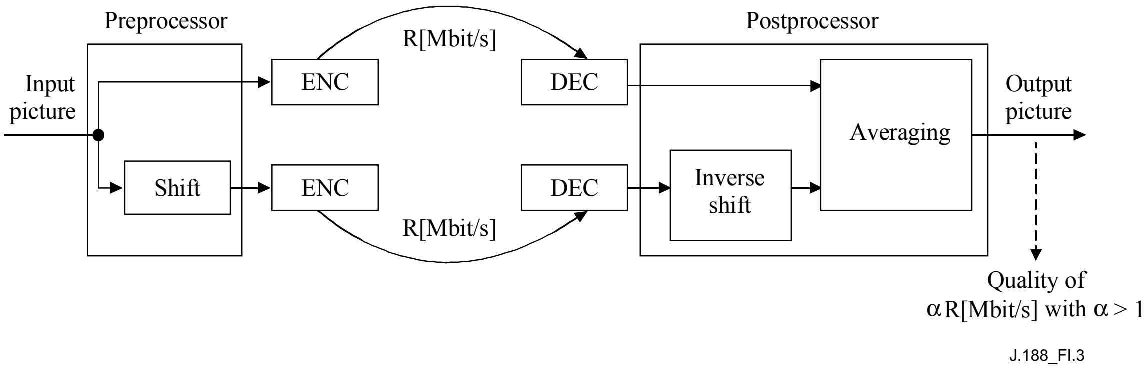

Figure I.3/J.188 – Configuration of efficient parallel transmission by use of spatial shifting

I.2.2 Temporal shifting method

MPEG-2 (ITU-T Rec. H.262), which is a major compression coding method for video transmission, adopts the GOP structure where motion-compensated prediction pictures and non-prediction pictures (I-pictures) exist periodically. Different picture types lead to different coding noise. If the timing of an I-picture is different between the two links, therefore, the coding noise is different between the two links even without the above-mentioned spatial shifting. We call this temporal shifting. Usually the encoders in parallel links are operated independently from each other, realizing temporal shifting. Therefore, no particular processing is needed before transmission. Thus in this case, simple averaging without pre- or post-processing for shifting is sufficient, as depicted in Figure I.4.

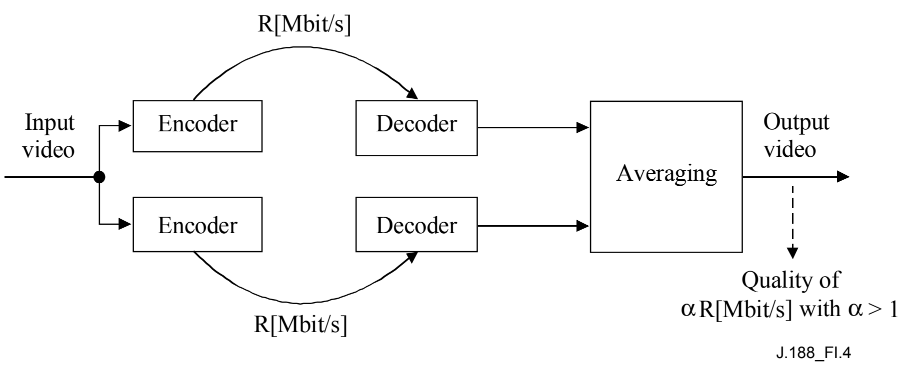

Figure I.4/J.188 – Configuration of efficient parallel transmission by use of temporal shifting

Appendix II

Failure detection and identification

II.1 Detection of failure

The picture is separated into many blocks, and the picture quality of both links is compared block by block. When the difference is larger than a threshold, a decision that a failure has occurred is issued. At this stage, however, it is still unknown which of the links has a failure. Figure II.1 shows the procedure.

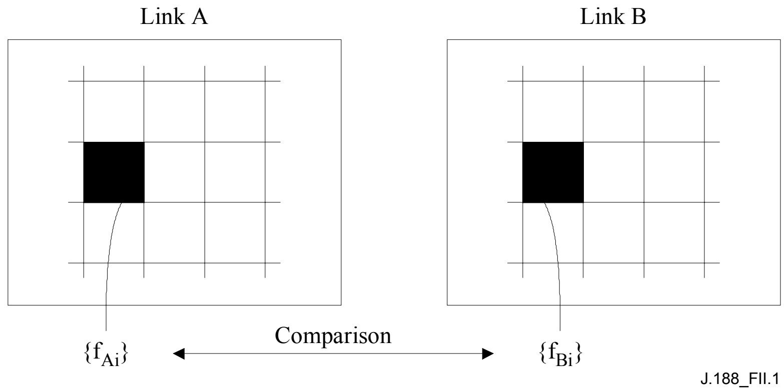

Figure II.1/J.188 – Comparison of image features

1) The processing is performed block by block in the typical size of 16 × 16. Let ***p*** denote the block position.

2) Calculate $\{f_{Ai}(p)\}, \{f_{Bi}(p)\}$ $(i = 1, ..., N)$, the image features in each block of two received pictures A and B. Here, N denotes the number of utilized image features.

   Compare the image features of A and B block by block, and when the difference for any one of the features is larger than a threshold at the block, i.e.:

$$\exists i \rightarrow |f_{Ai}(p) - f_{Bi}(p)| \rangle Th_i$$

   judge that one of the received pictures is corrupted by transmission failure at the block. Memorize this normal/corrupted information block by block. This information is utilized for identification of the disrupted link described below.

II.2 Identification of disrupted link

Next, the disrupted link is identified. By utilizing one characteristic of failure in digital video transmission, namely that both normal and corrupted parts coexist in a picture, a decision is made that the picture that contains a larger difference of quality between the two parts has a failure, and the disrupted link is thereby identified. Here, the corrupted part in a picture is known from the first stage. Figures II.2 and II.3 show the procedure.

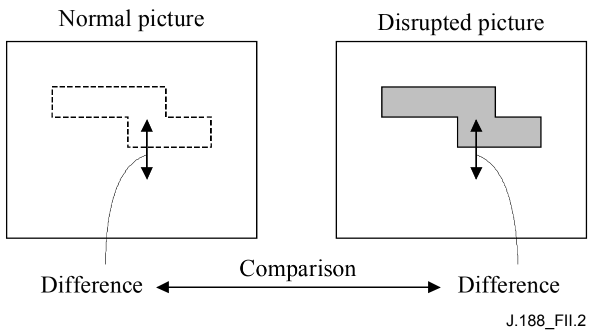

Figure II.2/J.188 – Comparison of quality difference

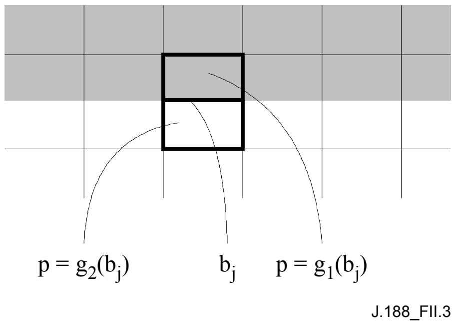

Figure II.3/J.188 – Boundary and blocks

1) Identification is performed field by field.

2) From the previous failure detection processing, we know which parts of the picture are corrupted. At this stage, however, it is still unknown which of the links has a failure.

3) Calculate $\boldsymbol{D} = \{D_i\}(i = 1, ..., N)$, the difference of image features between normal and corrupted areas, for each link. Specifically, calculate the sum of absolute difference along boundaries between normal and corrupted areas for each image feature as follows.

4) Let b_j (j = 1,...,$2n$) denote the boundary between blocks in the picture (a side of a block). Here n denotes the number of blocks in a picture (every block except those located on the edge of the picture is supposed to have two boundaries, namely right and bottom sides). Let, $\boldsymbol{g}_1(b_j)$, $\boldsymbol{g}_2(b_j)$ denote the two blocks which have b_j between them (Figure II.3). Calculate the difference of the ith image feature between the normal and corrupted areas in link A as follows. (Similar for link B).

$$D_{Ai} = \sum_{bj \in C} \left| f_{Ai}(g_1(b_j)) - f_{Ai}(g_2(b_j) \right|$$

Here, C represents all boundaries between normal and corrupted areas.

5) Compare $\|D_A\|$ and, $\|D_B\|$ and judge that link A has a failure when the former is larger, and vice versa. As $\|D\|$, use the maximum of the sums of absolute differences, i.e.:

$$\|D\| = \frac{\max \mathrm{D}_i}{1 \le i \le N}$$

Future J-series recommendations may also be utilized to identify the disrupted link.

Appendix III

Coding noise assessment

III.1 Theoretical background

The assessment of picture quality can be realized also utilizing the fact that coding noise in the two links is different (see Figure III.1).

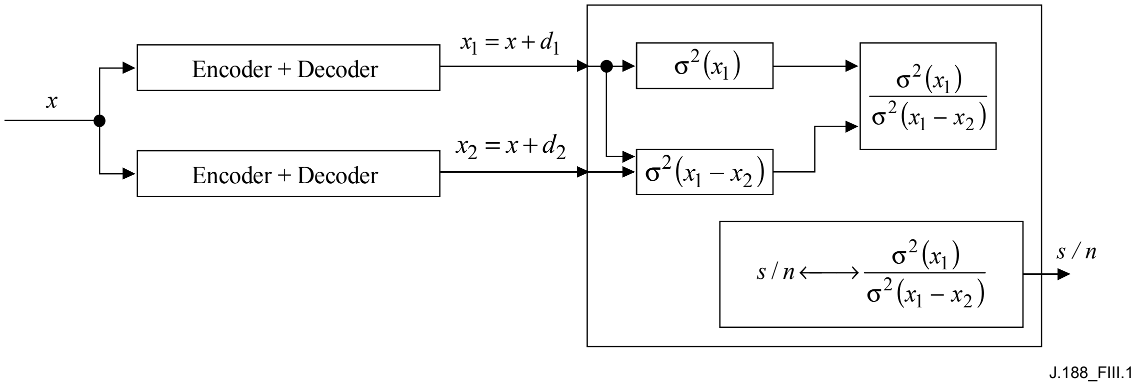

Figure III.1/J.188 – Assessment of SN ratio

Let x be the original signal, x_1, x_2 the two decoded signals, and d_1, d_2 their coding noise, respectively. s represents the variance of x, and n the power of d_1 and d_2. Let ρ_{xd1} be the correlation coefficient of x and d_1, and ρ_{d1d2} that of d_1 and d_2. The variance of x_1 is:

$$\sigma^2(x_1) = \sigma^2(x + d_1) = s + n + 2\sqrt{sn\rho_{xd1}} \tag{1}$$

The variance of the difference of x_1 and x_2 is:

$$\sigma^2(x_1 - x_2) = \sigma^2(d_1 - d_2) = 2n(1 - \rho_{d1d2}) \tag{2}$$

Then:

$$\frac{\sigma^2(x_1)}{\sigma^2(x_1 - x_2)} = \frac{1 + \frac{n}{s} + 2\rho_{xd1}\sqrt{\frac{n}{s}}}{2 - 2\rho_{d1d2}} \times \frac{s}{n} \tag{3}$$

If we assume that the coding noise is sufficiently smaller than the original signal and its correlation is also sufficiently small, i.e. $1 >> \frac{n}{s} + 2\rho_{xd1}\sqrt{\frac{n}{s}}$, we obtain:

$$10\log\frac{s}{n} = \alpha + 10\log\frac{\sigma^2(x_1)}{\sigma^2(x_1 - x_2)} \tag{4}$$

where $\alpha = 10\log(2 - 2\rho_{d1d2})$.

Experiments have shown that we can assume that α in Equation (4) depends only on the configuration of the parallel system, not on the kind of pictures. It is possible, therefore, to estimate the SNR of the received signal compared with the original by using only x_1 and x_2 that are obtained on the receiver side.

III.2 Method for SNR estimation

Figure III.1 shows the method.

1) In preparatory experiments, several kinds of pictures are encoded at several bitrates in the actual transmission scheme, and the relationship between *s*/*n* and

$$\frac{\sigma^2(x_1)}{\sigma^2(x_1 - x_2)}$$

is stored as a table.

2) During ordinary transmission,

$$\frac{\sigma^2(x_1)}{\sigma^2(x_1 - x_2)}$$

is measured from two decoded pictures, and then *s*/*n* is estimated by referring to the table.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	**Cable networks and transmission of television, sound programme and other multimedia signals**
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure and Internet protocol aspects
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2002