INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.147
(07/2002)

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

Measurement of the quality of service

Objective picture quality measurement method by use of in-service test signals

ITU-T Recommendation J.147

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.147

Objective picture quality measurement method by use of in-service test signals

Summary

This Recommendation presents an objective picture quality measurement method by use of in-service test signals in a digital television network. This method is separated into three procedures:

a) test signal generation at the transmission side using the invisible marker method;

b) test signal detection at measurement points; and

c) picture quality measurement using the result of test signal detection. A practical implementation is described in Appendix I, whilst the parameters for generating in-service test signals and the performance of this method are described in Appendices II and III, respectively.

Source

ITU-T Recommendation J.147 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 not 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

Introduction

The Reduced Reference (RR) and No Reference (NR) methods, which are currently being studied by an independent body, are expected to be covered in a future Recommendation for an objective measurement method of perceived picture quality that does not require reference pictures.

This Recommendation recommends a different framework from RR or NR for measuring picture quality without reference pictures by use of in-service test signals. In-service test signals are embedded into the active picture area of television signals by using the invisible marker method. Television signals with test signals are transmitted to the receiver side and the test signals are detected from the received pictures. Degradation of the test signals is measured to estimate picture quality degradation. This framework does not require an additional circuit for reference information because the reference information is already embedded into the transmitted pictures.

It may be possible to apply this method also to audio signals to ensure that the correct audio and video is matched within a multichannel environment, to ensure that audio and video are properly synchronized, and to measure any asynchronies, which are beyond the scope of this Recommendation.

ITU-T Recommendation J.147

Objective picture quality measurement method by use of in-service test signals

1 Scope

This Recommendation presents an objective picture quality measurement framework that works as an automatic quality monitoring system in a digital television network by use of in-service test signals. In-service test signals are applied for applications such as video transmission between different two locations. This method is applicable for either compressed or uncompressed transmission; however, this method is practical especially when lossy compression is applied in the transmission chain, because compression degradation masks degradation caused by test signals themselves. When PSNR is to be estimated by the method, it is a good idea to adapt the intensity of the test signals to the compression bit rate (see Appendix III). This method is applicable, for example, to transmission between studios and to broadcasting to end receivers.

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 references

– 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-R Recommendation BT.656-4 (1998), *Interfaces for digital component video signals in 525-line and 625-line television systems operating at the 4:2:2 level of Recommendation ITU-R BT.601 (Part A)*.

3 Terms, definitions and acronyms

This Recommendation uses the following abbreviations:

DEC	Decoder
ENC	Encoder
FDR	False Detection Rate
PSNR	Peak Signal-to-Noise Ratio
WHT	Walsh-Hadamard Transform
XOR	Exclusive OR

4 User requirements

User requirements for perceptual measurement methods of picture quality are given in ITU-T Rec. J.143.

5 Recommended framework

Configuration of the assumed transmission chain is shown in Figure 1. Figure 1 shows cascading connections of the transmission equipment (ENC and DEC). Either the compressed or uncompressed transmission can be applied for each single connection. A combination of compressed and uncompressed transmission connections can also be applied.

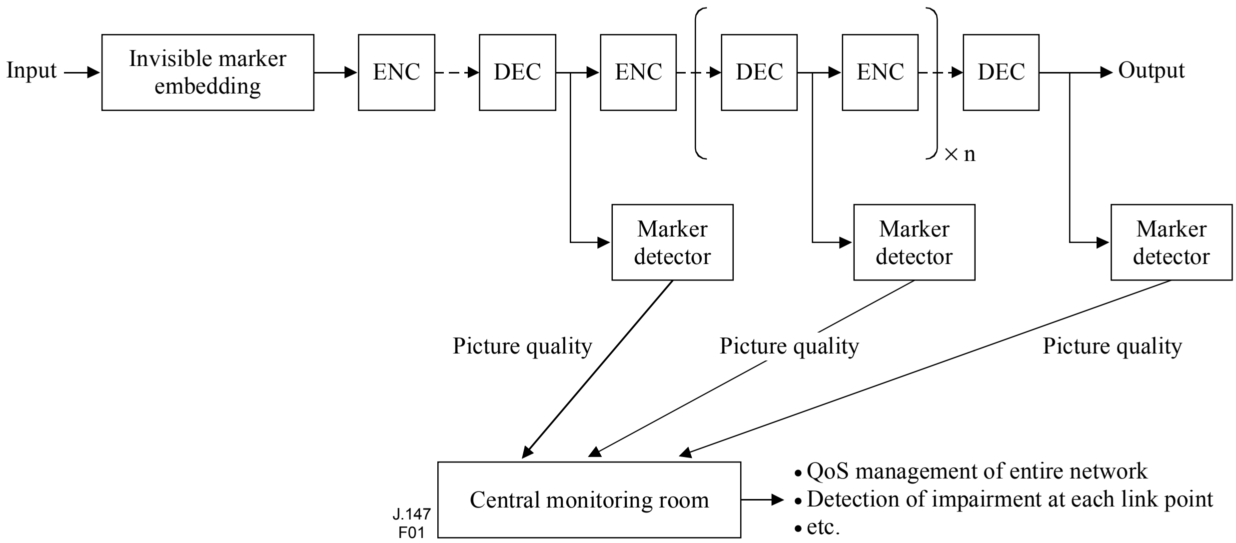

Figure 1/J.147 – Configuration of transmission chain

In-service test signals are embedded into input pictures and are detected at each link point of the cascading connection. Objective picture quality assessment score at each measurement point are transmitted to the central monitoring room via a dedicated circuit for remote monitoring which is separately constructed from video transmission. Since the picture quality of each link point can be monitored at the central monitoring room, network operators can recognize at which point the picture quality degradation (i.e. transmission trouble) has occurred.

In order to monitor picture quality by this configuration, the following items are required:

– embedding in-service test signal into source pictures;

– test signals detection from received pictures;

– picture quality assessment from the result of marker detection.

Appendix I describes a practical implementation.

Appendix I

Description of implementation method[1]

I.1 Marker embedding

Invisible markers are embedded into the active picture area of the video signal. Ancillary data of the video signal (e.g. ITU-R Rec. BT.656-4) should be transmitted transparently.

Marker embedding is performed at each pixel block. The selection of which and how many blocks the marker is embedded into is arbitrary. However, it is recommended to embed the markers into all the blocks in a frame/field when the measurement of an entire frame/field degradation is required.

The procedure for marker embedding is shown in Figure I.1. $x_B(n)$ denotes the input signal separated into pixel blocks. First, the spectrum of the signal $x_B(n)$ is spread by multiplying a PN sequence and then transform coefficients $X_{SS}[s]$ are derived by orthogonal transform. Invisible markers are embedded by choosing arbitrary frequency component s_i and manipulating those amplitude component A_i according to embedded bit *embedded bit*$_i$ and marker intensity M_i. The relationship between A_i, *embedded bit*$_i$ and M_i is shown in Tables I.1 and I.2.

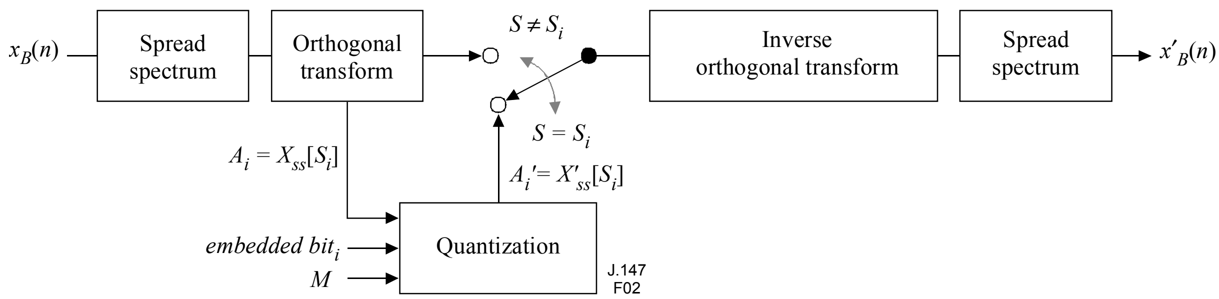

Figure I.1/J.147 – Marker embedding

Table I.1/J.147 – Relationship between A_i, *embedded bit$_i$* and M_i

	embedded bit$_i$ = 0	**embedded bit$_i$ = 1**
$\text{int}(A_i / M) = 2m$	$\left(2m + \frac{1}{2}\right)M$	$\left(2m - \frac{1}{2}\right)M$ if $A_i < \left(2m + \frac{1}{2}\right)M$
		$\left(2m + \frac{3}{2}\right)M$ otherwise
$\text{int}(A_i / M) = 2m + 1$	$\left(2m + \frac{1}{2}\right)M$ if $A_i < \left(2m + \frac{3}{2}\right)M$	$\left(2m + \frac{3}{2}\right)M$
	$\left(2m + \frac{5}{2}\right)M$ otherwise	

[1] This is one practical implementation. Alternative implementations are possible.

Table I.2/J.147 – Relationship between A_i, *embedded bit$_i$* and M_i

<table>
<tr><th></th><th>embedded bit$_i$ = 0</th><th colspan="2">embedded bit$_i$ = 1</th></tr>
<tr><td rowspan="2">$\text{round}(A_i / M) = 2m$</td><td rowspan="2">$2mM$</td><td>$(2m+1)M$</td><td>if $A_i < 2mM$</td></tr>
<tr><td>$(2m-1)M$</td><td>otherwise</td></tr>
<tr><td rowspan="2">$\text{round}(A_i / M) = 2m+1$</td><td>$(2m+2)M$ if $A_i < (2m+1)M$</td><td colspan="2" rowspan="2">$(2m+1)M$</td></tr>
<tr><td>$2mM$ otherwise</td></tr>
</table>

In Tables I.1 and I.2, m denotes an arbitrary integer and the value of *embedded bit$_i$* and M_i are also arbitrary. int(x) is the function that truncates the real value x and round(x) is the function that rounds up or down the real value x. The difference between function int(x) and round(x) is shown in Figures I.2 and I.3. Either of Table I.1 and Table I.2 can be applied for marker embedding, but it is recommended to apply Table I.1 when the amplitude A_i is always positive, i.e. $A_i > 0$. On the other hand, if A_i may be negative, the application of Table I.2 is recommended.

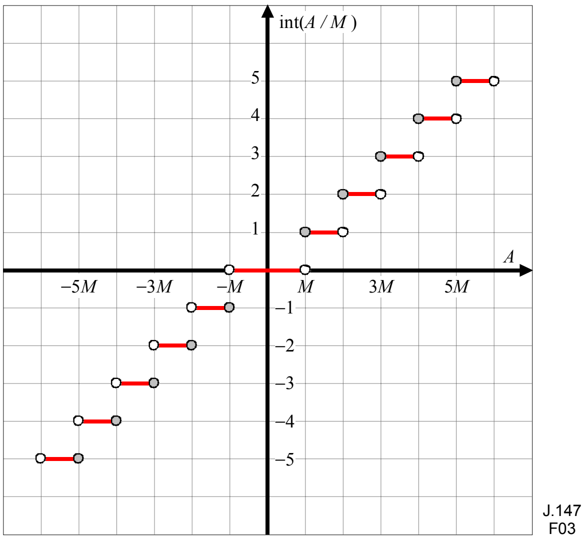

Figure I.2/J.147 – Definition of int(x)

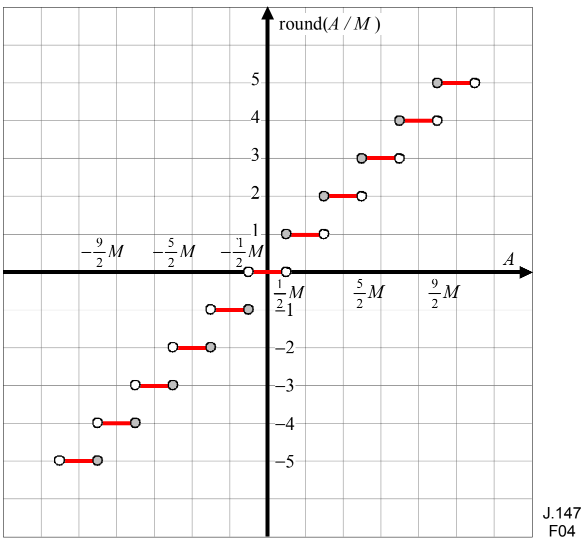

Figure I.3/J.147 – Definition of round(x)

When more than one markers are embedded, i.e. $N_m > 1$, a different value of the *embedded bit*$_i$ and M_i can be applied for each component i. After manipulating the transform coefficient by the above procedures, inverse orthogonal transform and inverse spread spectrum are applied and spatial domain signal $x'_B(n)$ is finally derived.

I.2 Marker detection

Invisible markers are detected at the receiver side and each link point of the transmission chain (see Figure I.4). In the detection side, either of the following formulae are applied to A_i'' to detect an embedded bit:

$$\begin{aligned}&\mathrm{int}(A_i''/M) = \mathit{even} \rightarrow \mathit{detected\ bit}_i = 0\\&\mathrm{int}(A_i''/M) = \mathit{odd} \rightarrow \mathit{detected\ bit}_i = 1\end{aligned} \qquad \text{if Table I.1 is applied for embedding}$$

$$\begin{aligned}&\mathrm{round}(A_i''/M) = \mathit{even} \rightarrow \mathit{detected\ bit}_i = 0\\&\mathrm{round}(A_i''/M) = \mathit{odd} \rightarrow \mathit{detected\ bit}_i = 1\end{aligned} \qquad \text{if Table I.2 is applied for embedding}$$

where A_i'' denotes the amplitude components of the signal of the received picture after spread spectrum. After detecting the binary information *detected bit*$_i$ from the received picture, the *embedded bit*$_i$ and *detected bit*$_i$ are compared. When these two bits are different, i.e.:

$$\mathrm{XOR}\ (\mathit{embedded\ bit}_i, \mathit{detected\ bit}_i) = 1$$

where XOR (a,b) denotes exclusive-OR between binary symbol a and b, the block is assumed to be falsely detected. False detection expresses that the transform coefficient A_i'' has been changed more than the marker intensity M_i and it implies that both the marker signal and video signal are degraded by transmission. Thus, the degradation of the video signal can be detected based on the result of detection of the marker signals.

In order to compare embedded and detected bits, the s_i, M_i, *embedded bit*$_i$ and the information which block the marker is embedded should be shared between the transmission side and each link point. Furthermore, the PN sequence used for the spread spectrum should be common to all the measurement points.

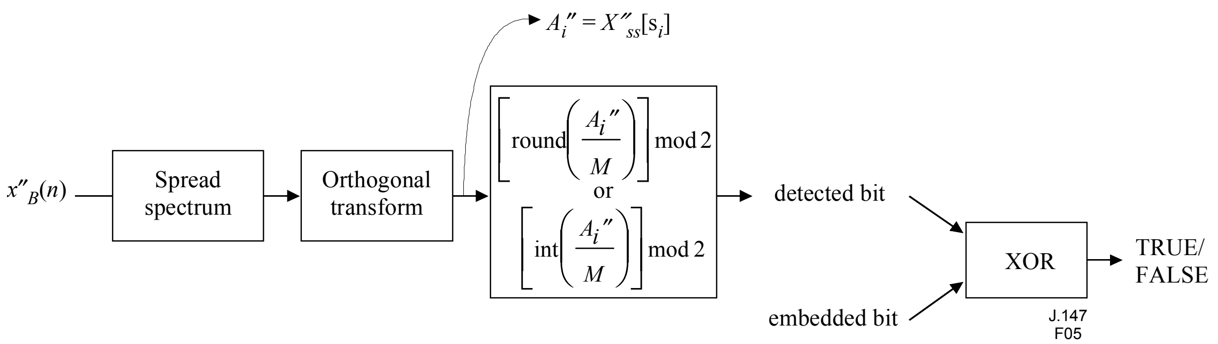

Figure I.4/J.147 – Marker detection

I.3 Picture quality measurement from the result of marker detection

Picture quality measurement of each frame/field is performed based on the result of marker detection. An arbitrary method can be applied to evaluate picture quality. In this clause, one example that uses the FDR (False Detection Rate) of the markers as a simple index for marker degradation is shown.

FDR is defined in the following equation:

$$\text{FDR} = \frac{\text{Number of false detected blocks in a frame (or field)}}{\text{Number of blocks in a frame}}$$

PSNR is one of the picture quality indexes that have a good correlation with the FDR. The estimated PSNR is expressed as follows:

$$\text{Estimated PSNR} = f(\text{FDR})$$

Where f denotes the correlation function. The above formula shows the estimated PSNR of one frame/field; however, the average PSNR of a certain period of frames/fields can be estimated when FDR is replaced with the averaged FDR.

It is possible to extend this method to detect problems with the transmission stream (e.g. packet headers, etc.)

Appendix II

Parameters for generating in-service test signals[2]

II.1 Marker intensity

Marker intensity M_i is an important parameter for embedding markers. M_i is called marker "intensity" because the power of the marker signal will increase according to M_i. As the marker signal will be a noise to the original pictures, it is desirable to keep its intensity as small as possible. However, if the marker intensity is not appropriate, it is possible that the marker signal is indistinguishable from the noise caused by transmission when the degradation caused by transmission is large. The marker intensity is thus determined in consideration of the configuration of the transmission chain, and the trade-off between the desired estimation accuracy and the degradation due to the markers.

II.2 Embedded bit

Arbitrary sequences can be used for embedded bits because they are used only for comparing with the detected bits. It is recommended that all 0 or all 1 sequences are used for the embedded bit to allow easy implementation. Since the marker signals are not tolerant against attacks (e.g. MPEG-2 compression) and do not ensure complete detection, embedding purposeful information is not suitable. It is recommended that the embedded information only be used for the purpose of picture quality measurements.

II.3 The number of embedded bits in a block

An arbitrary number of embedded bits N_m are available. Increasing N_m results in an increase of the number of embedded bits in a frame/field and therefore results in an increase of reference information. This leads to an accurate picture quality measurement; however, degradation of the original pictures also increases.

---

2 This refers to the particular implementation in Appendix I.

II.4 Blocksize

An arbitrary number of pixels can be chosen for the vertical and horizontal size of the pixel block in which the marker is embedded. However, for easy implementation, it is recommended to choose a size so that the fast calculation algorithm can be applied (e.g. 2^n pixels for Fast Fourier Transform (FFT)). For an accurate picture quality measurement, it is recommended to choose the same blocksize as that the compression-coding algorithm applies to when the compression is applied in the transmission chain.

II.5 PN sequences

PN sequence means a random sequence in which the elements consist of 1 or −1. For example, [1,1,–1,1,–1,–1,1,1,–1,–1,–1,1,1,–1,1,1,–1,...]. An arbitrary method can be applied to generate the PN sequence; however, the original pictures should be spread as randomly as possible for an accurate picture quality evaluation. A PN sequence of a longer period, at least longer than the number of pixels in the block, is therefore required.

II.6 Orthogonal transform and manipulation of its coefficients

As for orthogonal transform, an arbitrary method can be applied, such as DCT (Discrete Cosine Transform) and WHT (Walsh-Hadamard Transform). Note that the manipulation is performed to amplitude components. When the transform coefficients have phase components like DFT (Discrete Fourier Transform) coefficients, only amplitude components have to be manipulated while the phase components remain unchanged.

II.7 How to estimate picture quality from FDR

The relationship between FDR and picture quality is different from the configuration of the transmission chain. When MPEG-2 is applied for compression, the following relationship is obtained:

$$MSE \approx 2\left[\frac{\ln(FDR)}{M_i}\right]^2$$

Therefore, the relationship between FDR and PSNR is derived as follows:

$$PSNR = a\log(\ln(FDR)) + b$$

where a and b denote the real constants which are variable to the implementation of codec, especially quantization specification. These constants can be obtained by preliminary experiment using the reference pictures.

Appendix III

Performance evaluation[3]

III.1 Experiment conditions

Six test sequences with invisible markers are encoded by MPEG-2 TM5 codec and the Peak Signal-to-Noise Ratio (PSNR) and False Detection Rate (FDR) are measured by the decoded pictures to obtain the FDR-PSNR relationship. The degradation of the original picture is also examined.

---

[3] The result in this appendix indicates the measurement made for the implementation shown in Appendix I. Independent results to show the relationship with the subjective scores are in process.

Table III.1 shows the simulation conditions. The proposed method can be applied to the estimation of both the luminance and chrominance component; however, in this experiment the markers are embedded only into the luminance component and the PSNR of the luminance component is estimated.

Table III.1/J.147 – Experiment conditions

Test sequences (Note)	Cheerleaders, Flamingoes, Green Leaves, Marching in, Mobile and Calendar, Soccer Action
Picture size	704(H) × 480(V)
Format	4:2:2, 30 fps
Codec	MPEG-2 Test Model 5
Bit rate	a) 10, 20, 30, 40 Mbit/s b) 6, 8, 10, 12 Mbit/s
Picture quality measure	PSNR
Blocksize	8 × 8
Market intensity	a) M = 60 b) M = 100
NOTE – See ITU-R Rec. BT.1210-2.	

The video format is 4:2:2, 8 bits/pixel, 30 fps, which is conformable to ITU-R Rec. BT.601. The markers are embedded into 8 × 8 pixel blocks and only one marker is embedded into one block. As for the bit rates, two sets of bit rates are examined:

a) 10 to 40 Mbit/s assuming contribution;

b) 6 to 12 Mbit/s assuming primary distribution.

A different marker intensity is selected for each bit rate set.

III.2 Experiment Result

Figures III.1 and III.2 show the FDR-PSNR relationship of bit rate set a) and b), respectively. This relationship cannot be obtained if the reference picture is not present; however, once this relationship is obtained, the PSNR can be estimated from the FDR since this relationship is independent of bit rates and picture contents.

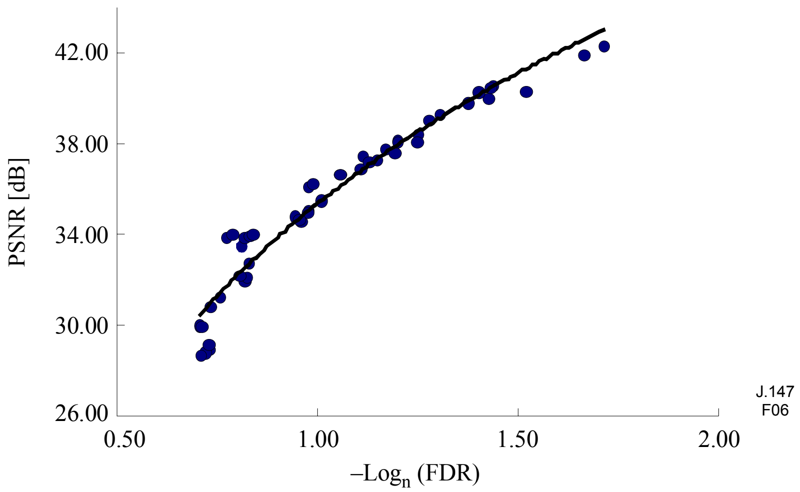

Figure III.1/J.147 – FDR-PSNR relationship (bit rate set a))

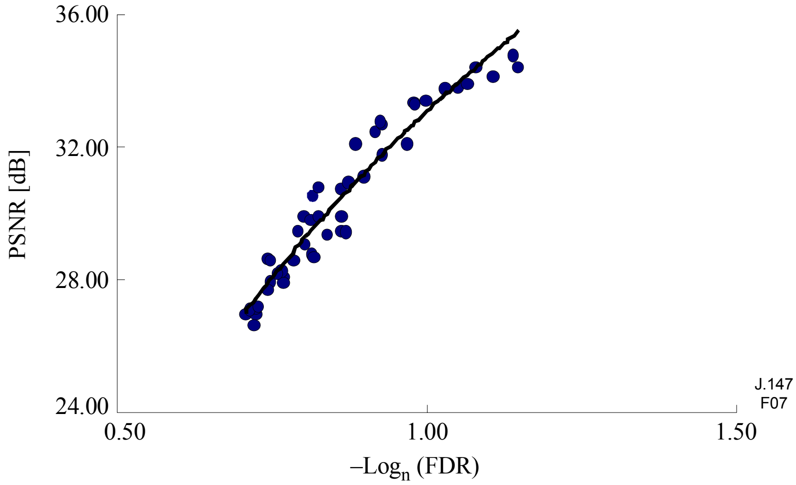

Figure III.2/J.147 – FDR-PSNR relationship (bit rate set b))

III.2.1 Estimation accuracy

Estimation accuracy is defined as the distance of the plotted point and the regression curve. Table III.2 shows the average estimation accuracy and the variance of the estimation error for each bit rate set. From this table, the estimation accuracy is about 0.5 to 0.6 dB, confirming that highly accurate PSNR estimation is possible using this method.

Table III.2/J.147 – Average estimation accuracy and variance

	Estimation accuracy [dB]	Variance [dB]
Simulation a)	0.59	0.74
Simulation b)	0.50	0.37

III.2.2 Degradation of the original picture caused by marker embedding

Table III.3 shows the PSNR of each picture title after marker embedding. Since the marker embedding causes a manipulation of the original signal, the marker signal will be a noise signal for the original picture. However, the power of the marker signal is so small that the PSNR of the marker embedded picture is over 49 dB as shown in Table III.3. This is well over 40 dB and not a perceivable level, and thus the marker signal cannot be seen by human eye. For reference, the typical PSNR of contribution and primary distribution is 30 to 40 dB, as shown in Figure III.1. In addition, the additional required bandwidth due to embedding markers has been found to be negligibly small with this marker intensity.

Table III.3/J.147 – PSNR of the original picture after marker embedding

Test sequence name	Simulation a)	Simulation b)
Cheerleaders	51.46	49.19
Flamingoes	51.45	49.26
Green Leaves	51.45	49.12
Marching in	51.44	49.13
Mobile and Calendar	51.45	49.15
Soccer Action	51.47	49.10

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