



**Recommendation ITU-R BT.1885**  
(03/2011)

**Objective perceptual video quality  
measurement techniques for standard  
definition digital broadcast television  
in the presence of a reduced  
bandwidth reference**

**BT Series**  
**Broadcasting service**  
**(television)**



International  
Telecommunication  
Union

**Documents**

Document Sharing Hub

## Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

## Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from <http://www.itu.int/ITU-R/go/patents/en> where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

### Series of ITU-R Recommendations

(Also available online at <http://www.itu.int/publ/R-REC/en>)

| Series     | Title                                                                                |
|------------|--------------------------------------------------------------------------------------|
| <b>BO</b>  | Satellite delivery                                                                   |
| <b>BR</b>  | Recording for production, archival and play-out; film for television                 |
| <b>BS</b>  | Broadcasting service (sound)                                                         |
| <b>BT</b>  | <b>Broadcasting service (television)</b>                                             |
| <b>F</b>   | Fixed service                                                                        |
| <b>M</b>   | Mobile, radiodetermination, amateur and related satellite services                   |
| <b>P</b>   | Radiowave propagation                                                                |
| <b>RA</b>  | Radio astronomy                                                                      |
| <b>RS</b>  | Remote sensing systems                                                               |
| <b>S</b>   | Fixed-satellite service                                                              |
| <b>SA</b>  | Space applications and meteorology                                                   |
| <b>SF</b>  | Frequency sharing and coordination between fixed-satellite and fixed service systems |
| <b>SM</b>  | Spectrum management                                                                  |
| <b>SNG</b> | Satellite news gathering                                                             |
| <b>TF</b>  | Time signals and frequency standards emissions                                       |
| <b>V</b>   | Vocabulary and related subjects                                                      |

**Note:** This ITU-R Recommendation was approved in English under the procedure detailed in Resolution ITU-R 1.

Electronic Publication  
Geneva, 2011

© ITU 2011

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without written permission of ITU.

## RECOMMENDATION ITU-R BT.1885

**Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a reduced bandwidth reference**

(2011)

**Scope**

This Recommendation describes objective video quality assessment methods for standard definition digital broadcast television which can measure the perceptual video quality of mobile and stationary receiving conditions when the features extracted from the reference video signal are readily available at the measurement point.

The ITU Radiocommunication Assembly,

*considering*

- a) that the ability to measure automatically the impairments of broadcast video signals has been desirable;
- b) that the perceptual video quality of mobile reception may dynamically change depending on reception conditions;
- c) that objective measurement of perceived video quality may usefully complement subjective assessment methods;
- d) that three video quality measurement techniques for standard definition digital broadcast television in the presence of a reduced reference have been proposed to ITU-R and they have been found to provide equivalent, well-aligned results;
- e) that objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference have been specified in Recommendation ITU-R BT.1683,

*recommends*

- 1 that any of the objective video quality models given in Annex 1 should be used for objective measurement of perceived video quality for standard definition digital broadcast television in the presence of a reduced bandwidth reference.

**1 Introduction**

This RRNR-TV test addresses images as defined in Recommendation ITU-R BT.601-6 and two types of models: reduced reference (RR), and no reference (NR). RR models have limited bandwidth access to the source video, and NR models do not have access to the source video.

The HRCs in each experiment encompassed both coding only artefacts and coding with transmission errors. The coding schemes examined were MPEG-2 and H.264 (MPEG-4 part 10). The MPEG-2 coders were run at a variety of bit rates from 1.0 to 5.5 Mbit/s. The H.264 coders were run at a variety of bit rates ranging from 1.0 to 3.98 Mbit/s. Each experiment included 12 source sequences, of which two were secret source. Each experiment included 34 HRCs, and 156 processed video sequences (PVSSs). Of these PVSSs, 40 contained transmission errors and 116 contained coding only.

## 1.1 Application

This Recommendation provides video quality estimations for video classes TV3 to MM5B, as defined in ITU-T Recommendation P.911, Annex B. The applications for the estimation models described in this Recommendation include but are not limited to:

1. potentially real-time, in-service quality monitoring at the source;
2. remote destination quality monitoring when side-channels are available for features extracted from source video sequences;
3. quality measurement for monitoring of a storage or transmission system that utilizes video compression and decompression techniques, either a single pass or a concatenation of such techniques;
4. lab testing of video systems.

## 1.2 Limitations

The estimation models described in this Recommendation cannot be used to fully replace subjective testing. Correlation values between two carefully designed and executed subjective tests (i.e. in two different laboratories) normally fall within the range 0.95 to 0.98. If this Recommendation is utilized to compare different codecs, it is advisable to use a quantitative method (such as that in ITU-T Recommendation J.149) to determine the model's accuracy for that particular context.

The models in this Recommendation were validated by measuring objective video quality that exhibits frame freezes up to 2 s.

The models in this Recommendation were not validated for measuring objective video quality that has a steadily increasing delay (e.g. video which does not discard missing frames after a frame freeze).

It should be noted that in case of new coding and transmission technologies producing artefacts which were not included in this evaluation, the objective evaluation models may produce erroneous results. Here a subjective evaluation is required.

## 2 References

The following ITU 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

Recommendation ITU-R BT.500-12 – Methodology for the subjective assessment of the quality of television pictures.

ITU-T Recommendation P.910 (2008) – Subjective video quality assessment methods for multimedia applications.

ITU-T Recommendation P.911 (1998) – Subjective audiovisual quality assessment methods for multimedia applications.

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

ITU-T Recommendation J.244 (2008) – Full reference and reduced reference calibration methods for video transmission systems with constant misalignment of spatial and temporal domains with constant gain and offset.

## 2.2 Informative references

ITU-T Recommendation J.149 (1998) – Subjective audiovisual quality assessment methods for multimedia applications.

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 P.931 (1998) – Multimedia communications delay, synchronization and frame rate measurement.

ITU-T Recommendation J.148 (2003) – Requirements for an objective perceptual multimedia quality model.

ITU-T Recommendation H.261 (1993) – Video codec for audiovisual services at p x 64 kbits.

ITU-T Recommendation H.263 (1996) – Video coding for low bit rate communication.

ITU-T Recommendation H.263 (1998) – Video coding for low bit rate communication (H.263+).

ITU-T Recommendation H.264 (2003) – Advanced video coding for generic audiovisual services.

VQEG – Validation of reduced-reference and no-reference objective models for standard definition television, Phase I, 2009.

## 3 Definitions

### 3.1 Terms defined elsewhere:

This Recommendation uses the following terms defined elsewhere:

**3.1.1 subjective assessment (picture)** (ITU-T Recommendation J.144): optional quoted definition.

**3.1.2 objective perceptual measurement (picture)** (ITU-T Recommendation J.144): optional quoted definition.

**3.1.3 Proponent** (ITU-T Recommendation J.144): optional quoted definition.

### 3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

**3.2.1 Anomalous frame repetition:** is defined as an event where the HRC outputs a single frame repeatedly in response to an unusual or out of the ordinary event. Anomalous frame repetition includes but is not limited to the following types of events: an error in the transmission channel, a change in the delay through the transmission channel, limited computer resources impacting the decoder's performance, and limited computer resources impacting the display of the video signal.

**3.2.2 Constant frame skipping:** is defined as an event where the HRC outputs frames with updated content at an effective frame rate that is fixed and less than the source frame rate.

**3.2.3 Effective frame rate:** is defined as the number of unique frames (i.e. total frames – repeated frames) per second.

**3.2.4 Frame rate:** is defined as the number of unique frames (i.e. total frames – repeated frames) per second.

**3.2.5 Intended frame rate:** is defined as the number of video frames per second physically stored for some representation of a video sequence. The frame rate shall be constant. Two examples of constant intended frame rates are a BetacamSP® tape containing 25 fps and a VQEG FR-TV Phase I compliant 625-line YUV file containing 25 fps; these both have an intended frame rate of 25 fps.

**3.2.6 Live network conditions:** are defined as errors imposed upon the digital video bit stream as a result of live network conditions.

**3.2.7 Pausing with skipping:** is defined as events where the video pauses for some period of time and then restarts with some loss of video information. In pausing with skipping, the temporal delay through the system will vary about an average system delay, sometimes increasing and sometimes decreasing. One example of pausing with skipping is a pair of IP Videophones, where heavy network traffic causes the IP Videophone display to freeze briefly; when the IP Videophone display continues, some content has been lost. Constant frame skipping and variable frame skipping are subsets of pausing with skipping. A processed video sequence containing pausing with skipping will be approximately the same duration as the associated original video sequence.

**3.2.8 Pausing without skipping:** is defined as any event where the video pauses for some period of time and then restarts without losing any video information. Hence, the temporal delay through the system must increase.

**3.2.9 Refresh rate:** is defined as the rate at which the display is updated.

**3.2.10 Simulated transmission errors:** are defined as errors imposed upon the digital video bit stream in a highly controlled environment. Examples include simulated packet loss rates and simulated bit errors.

**3.2.11 Source frame rate (SFR):** is the intended frame rate of the original source video sequences. The source frame rate is constant. For the VQEG RRNR-TV test the SFR was either 25 fps or 30 fps.

**3.2.12 Transmission errors:** are defined as any error imposed on the video transmission. Example types of errors include simulated transmission errors and live network conditions.

**3.2.13 Variable frame skipping:** is defined as an event where the HRC outputs frames with updated content at an effective frame rate that changes with time. The temporal delay through the system will increase and decrease with time, varying about an average system delay. A processed video sequence containing variable frame skipping will be approximately the same duration as the associated original video sequence.

## 4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

|        |                                                                                 |
|--------|---------------------------------------------------------------------------------|
| ACR    | Absolute category rating (see ITU-T Recommendation P.910)                       |
| ACR-HR | Absolute category rating with hidden reference (see ITU-T Recommendation P.910) |
| AVI    | Audio video interleave                                                          |
| DMOS   | Difference mean opinion score                                                   |
| FR     | Full reference                                                                  |
| FRTV   | Full reference television                                                       |
| HRC    | Hypothetical reference circuit                                                  |

|      |                                     |
|------|-------------------------------------|
| NR   | No (or Zero) reference              |
| PSNR | Peak signal-to-noise ratio          |
| PVS  | Processed video sequence            |
| RMSE | Root mean square error              |
| RR   | Reduced reference                   |
| SFR  | Source frame rate                   |
| SRC  | Source reference channel or circuit |
| VQEG | Video Quality Experts Group         |
| YUV  | Colour space                        |

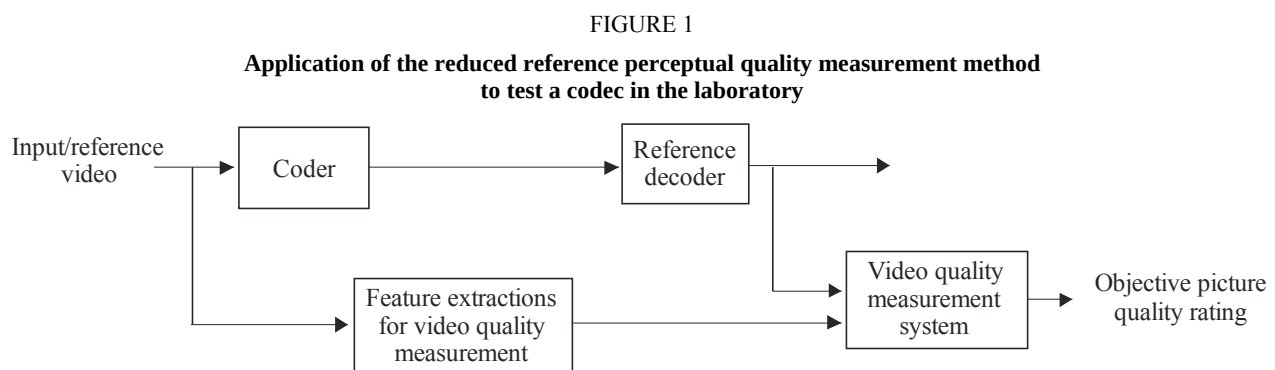
## 5 Conventions

None.

## 6 Description of the reduced reference measurement method

The double-ended measurement method with reduced reference, for objective measurement of perceptual video quality, evaluates the performance of systems by making a comparison between features extracted from the undistorted input, or reference, video signal at the input of the system, and the degraded signal at the output of the system (see Fig. 1).

Figure 1 shows an example of application of the reduced reference method to test a codec in the laboratory.



BT.1885-01

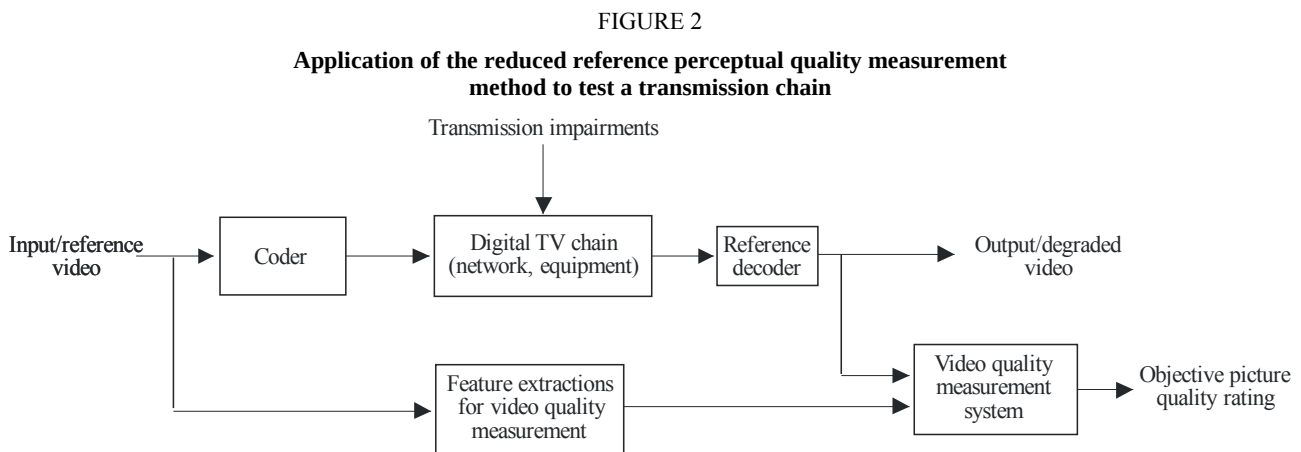
The comparison between input and output signals may require a temporal alignment or a spatial alignment process, the latter to compensate for any vertical or horizontal picture shifts or cropping. It also may require correction for any offsets or gain differences in both the luminance and the chrominance channels. The objective picture quality rating is then calculated, typically by applying a perceptual model of human vision.

Alignment and gain adjustment is known as normalization. This process is required because most reduced reference methods compare the features extracted from reference pictures and processed pictures on what is effectively a pixel-by-pixel basis. An example would be calculation of peak signal-to-noise ratio (PSNR). Only time-invariant static changes in the video are removed, dynamic changes due to system under test are measured as part of the quality rating calculation. ITU-T Recommendations J.244 and ITU-T J.144 provide standard methods for reporting values needed for

normalizing video prior to objective quality assessment. The video quality metrics described in Annex of this Recommendation include associated normalization methods. Alternate normalization methods can be used for the video quality metrics of Annex provided they deliver the required normalization accuracy.

As the video quality metrics are typically based on approximations to human visual responses, rather than on the measurement of specific coding artefacts, they are in principle equally valid for analogue systems and for digital systems. They are also in principle valid for chains where analogue and digital systems are mixed, or where digital compression systems are concatenated.

Figure 2 shows an example of the application of the reduced reference method to test a transmission chain.



BT.1885-02

In this case a reference decoder is fed from various points in the transmission chain, e.g. the decoder can be located at a point in the network, as in Fig. 2, or directly at the output of the encoder as in Fig. 1. If the digital transmission chain is transparent, the measurement of objective picture quality rating at the source is equal to the measurement at any subsequent point in the chain.

It is generally accepted that the full reference method provides the best accuracy for perceptual picture quality measurements. The method has been proven to have the potential for high correlation with subjective assessments made in conformity with the ACR-HR methods specified in ITU-T Recommendation P.910.

## 7 Findings of the Video Quality Experts Group

Studies of perceptual video quality measurements are conducted in an informal group, called Video Quality Experts Group (VQEG), which reports to ITU-T Study Groups 9 and 12 and Radiocommunications Study Group 6. The recently completed RRNR-TV test of VQEG assessed the performance of proposed reduced reference perceptual video quality measurement algorithms for ITU-R 601-6 image formats. formats.

Based on the present evidence, six RR methods (Model\_A 15k, Model\_A 80k, Model\_A 256k, Model\_C 80k, Model\_B 80k (525-line only), Model\_B 256k (525-line only)) can be recommended by ITU-T at this time.

The technical descriptions of these models can be found in Annexes A through C respectively. Note that the ordering of Annexes is purely arbitrary and provides no indication of quality prediction performance.

Tables 1 and 2 show significance tests in the VQEG RRNR-TV test. For the 525-line format, four models (Model\_A 15k, Model\_A 80k, Model\_A 256k, Model\_C 80k) are statistically better than PSNR and two models (Model\_B 80k, Model\_B 256k) are statistically equivalent to PSNR. It should be noted that PSNR was computed by NTIA using an exhaustive search of calibration limits. For the 625-line format, four models (Model\_A 15k, Model\_A 80k, Model\_A 256k, Model\_C 80k) are statistically equivalent and are statistically better than PSNR.

TABLE 1  
Significance test for the 525-line format

| 525-line format | Compare best | Compare PSNR | Correlation |
|-----------------|--------------|--------------|-------------|
| Model_A 15k     | 1            | 1            | 0.906       |
| Model_A 80k     | 1            | 1            | 0.903       |
| Model_A 256k    | 1            | 1            | 0.903       |
| Model_C 80k     | 1            | 1            | 0.882       |
| Model_B 80k     | 0            | 1            | 0.795       |
| Model_B 256k    | 0            | 1            | 0.803       |
| PSNR_NTIA       | 0            | 1            | 0.826       |

NOTE 1 – “1” in “Compare best” indicates that this model is statistically equivalent to the top performing model. “0” indicates that this model is not statistically equivalent to the top performing model. “1” in “Compare PSNR” indicates that this model is statistically equivalent to the top performing model. “0” indicates that this model is not statistically equivalent to the top performing model.

TABLE 2  
Significance test for the 625 format

| 625-line format | Compare best | Compare PSNR | Correlation |
|-----------------|--------------|--------------|-------------|
| Model_A 15k     | 1            | 1            | 0.894       |
| Model_A 80k     | 1            | 1            | 0.899       |
| Model_A 256k    | 1            | 1            | 0.898       |
| Model_C 80k     | 1            | 1            | 0.866       |
| PSNR_NTIA       | 0            | 1            | 0.857       |

NOTE 1 – “1” in “Compare best” indicates that this model is statistically equivalent to the top performing model. “0” indicates that this model is not statistically equivalent to the top performing model. “1” in “Compare PSNR” indicates that this model is statistically equivalent to the top performing model. “0” indicates that this model is not statistically equivalent to the top performing model.

Tables 3 and 4 provide informative details on the model’s performances in the VQEG RRNR-TV test.

TABLE 3

**Informative description on the model's performances in the VQEG RRNR-TV test (525-line format)**

| 525-line format     | Correlation | RMSE  | OR    |
|---------------------|-------------|-------|-------|
| <b>Model_A 15k</b>  | 0.906       | 0.418 | 0.385 |
| <b>Model_A 80k</b>  | 0.903       | 0.423 | 0.378 |
| <b>Model_A 256k</b> | 0.903       | 0.424 | 0.378 |
| <b>Model_B 80k</b>  | 0.795       | 0.598 | 0.667 |
| <b>Model_B 256k</b> | 0.803       | 0.587 | 0.647 |
| <b>Model_C 80k</b>  | 0.882       | 0.465 | 0.513 |
| <b>PSNR_NTIA</b>    | 0.826       | 0.556 | 0.571 |

TABLE 4

**Informative description on the model's performances in the VQEG RRNR-TV test (625-line format)**

| 625-line format     | Correlation | RMSE  | OR    |
|---------------------|-------------|-------|-------|
| <b>Model_A 15k</b>  | 0.894       | 0.524 | 0.468 |
| <b>Model_A 80k</b>  | 0.899       | 0.513 | 0.462 |
| <b>Model_A 256k</b> | 0.898       | 0.516 | 0.468 |
| <b>Model_C_80k</b>  | 0.866       | 0.585 | 0.583 |
| <b>PSNR_NTIA</b>    | 0.857       | 0.605 | 0.564 |

## Annex A

### Model A: Yonsei University reduced reference method

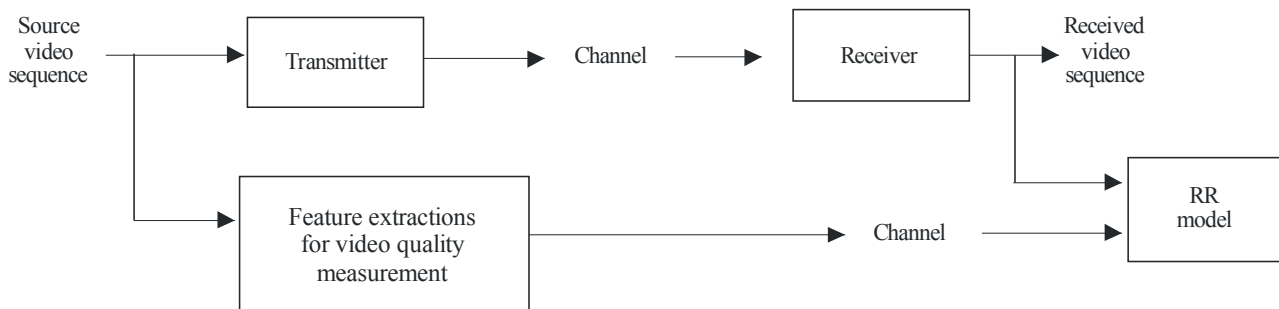
#### 1 Introduction

Although PSNR has been widely used as an objective video quality measure, it is also reported that it does not well represent perceptual video quality. By analysing how humans perceive video quality, it is observed that the human visual system is sensitive to degradation around the edges. In other words, when the edge pixels of a video are blurred, evaluators tend to give low scores to the video even though the PSNR is high. Based on this observation, the reduced reference models which mainly measure edge degradations have been developed.

Figure 3 illustrates how a reduced-reference model works. Features which will be used to measure video quality at a monitoring point are extracted from the source video sequence and transmitted. The Table 5 shows the side-channel bandwidths for the features, which have been tested in the VQEG RRNR-TV test.

FIGURE 3

Block diagram of reduced reference model



BT.1885-03

TABLE 5

Side-channel bandwidths

| Video format | Tested bandwidths                |
|--------------|----------------------------------|
| 525 format   | 15 kbit/s, 80 kbit/s, 256 kbit/s |
| 625 format   | 15 kbit/s, 80 kbit/s, 256 kbit/s |

## 2 The EPSNR reduced-reference models

### 2.1 Edge PSNR

The RR models mainly measure on edge degradations. In the models, an edge detection algorithm is first applied to the source video sequence to locate the edge pixels. Then, the degradation of those edge pixels is measured by computing the mean squared error. From this mean squared error, the edge PSNR (EPSNR) is computed.

One can use any edge detection algorithm, though there may be minor differences in the results. For example, one can use any gradient operator to locate edge pixels. A number of gradient operators have been proposed. In many edge detection algorithms, the horizontal gradient image  $g_{horizontal}(m,n)$  and the vertical gradient image  $g_{vertical}(m,n)$  are first computed using gradient operators. Then, the magnitude gradient image  $g(m,n)$  may be computed as follows:

$$g(m,n) = |g_{horizontal}(m,n)| + |g_{vertical}(m,n)|$$

Finally, a thresholding operation is applied to the magnitude gradient image  $g(m,n)$  to find edge pixels. In other words, pixels whose magnitude gradients exceed a threshold value are considered as edge pixels.

Figures 4 to Fig. 8 illustrate the procedure. Figure 4 shows a source image. Figure 5 shows a horizontal gradient image  $g_{horizontal}(m,n)$ , which is obtained by applying a horizontal gradient operator to the source image of Fig. 4. Figure 6 shows a vertical gradient image  $g_{vertical}(m,n)$ , which is obtained by applying a vertical gradient operator to the source image of Fig. 4. Figure 7 shows the magnitude gradient image (edge image) and Fig. 8 shows the binary edge image (mask image) obtained by applying thresholding to the magnitude gradient image of Fig. 7.

FIGURE 4

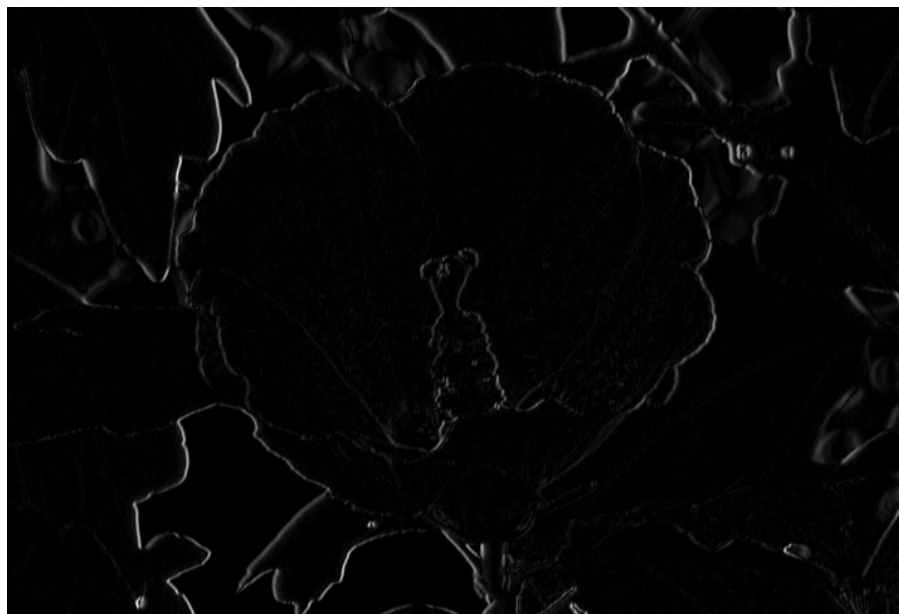
A source image (original image)



BT.1885-04

FIGURE 5

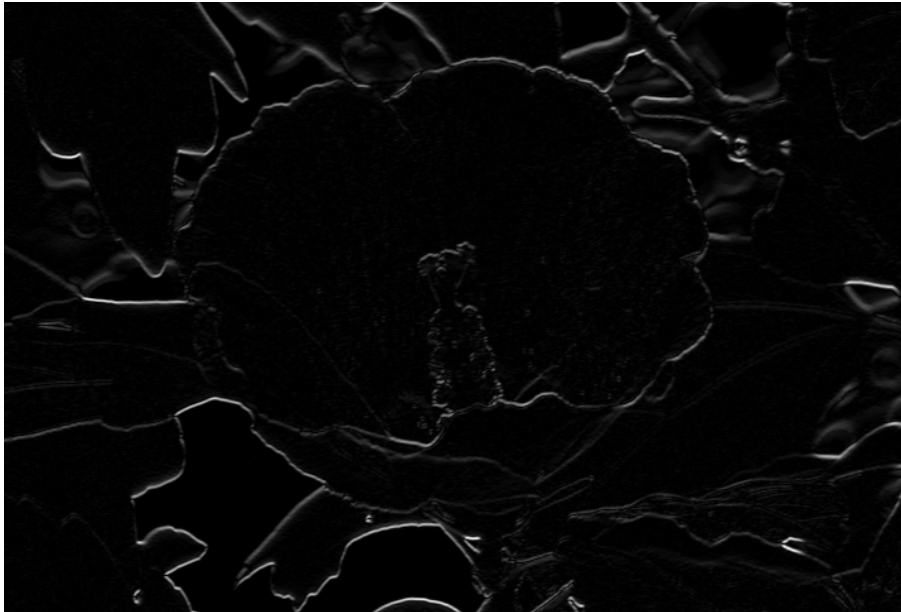
A horizontal gradient image, which is obtained by applying a horizontal gradient operator to the source image of Fig. 4



BT.1885-05

FIGURE 6

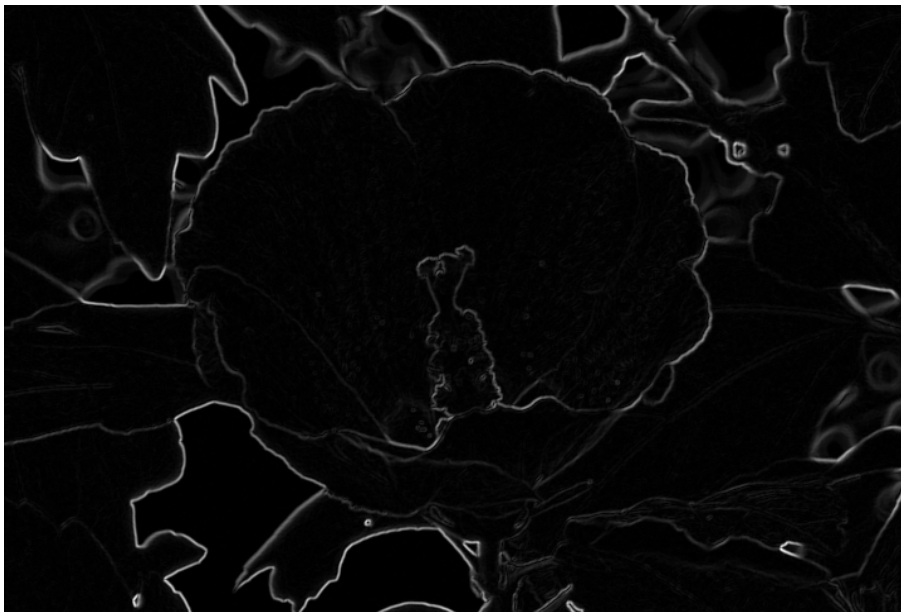
A vertical gradient image, which is obtained by applying a vertical gradient operator to the source image of Fig. 4



BT.1885-06

FIGURE 7

A magnitude gradient image

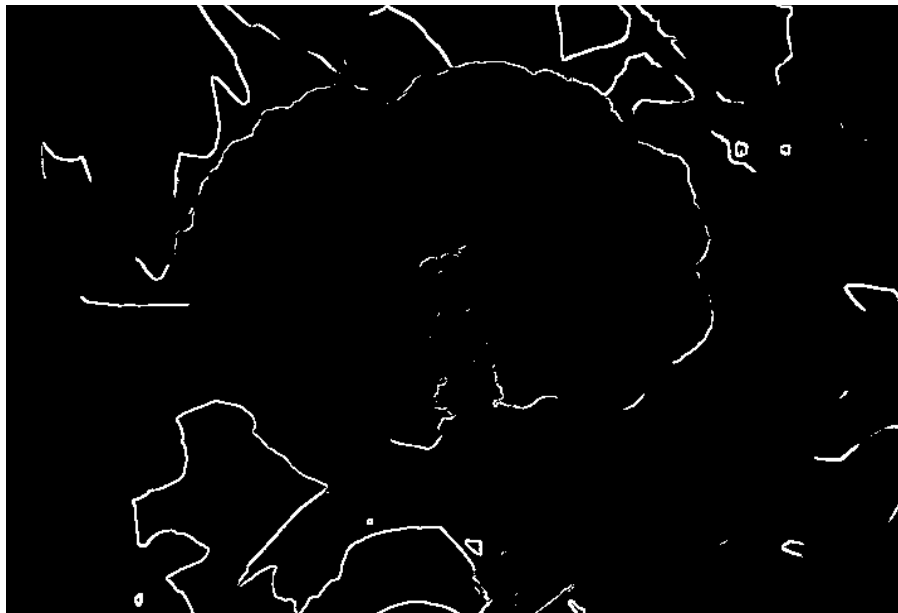


BT.1885-07

Copyright International Telecommunication Union/ITU Radiocommunication Sector

FIGURE 8

A binary edge image (mask image) obtained by applying thresholding to the magnitude gradient image of Fig. 7



BT.1885-08

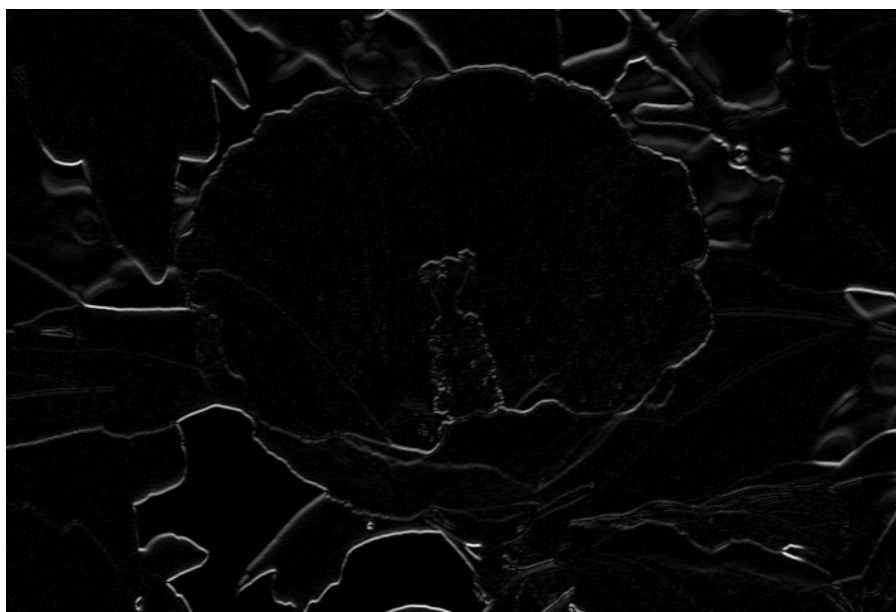
Alternatively, one may use a modified procedure to find edge pixels. For instance, one may first apply a vertical gradient operator to the source image, producing a vertical gradient image. Then, a horizontal gradient operator is applied to the vertical gradient image, producing a modified successive gradient image (horizontal and vertical gradient image). Finally, a thresholding operation may be applied to the modified successive gradient image to find edge pixels. In other words, pixels of the modified successive gradient image, which exceed a threshold value, are considered as edge pixels. Figures 9 to Fig. 12 illustrate the modified procedure. Figure 9 shows a vertical gradient image  $g_{vertical}(m,n)$ , which is obtained by applying a vertical gradient operator to the source image of Fig. 4. Figure 10 shows a modified successive gradient image (horizontal and vertical gradient image), which is obtained by applying a horizontal gradient operator to the vertical gradient image of Fig. 9. Figure 11 shows the binary edge image (mask image) obtained by applying thresholding to the modified successive gradient image of Fig. 10.

It is noted that both methods can be understood as an edge detection algorithm. One may choose any edge detection algorithm depending on the nature of videos and compression algorithms. However, some methods may outperform other methods.

Thus, in the model, an edge detection operator is first applied, producing edge images (see Figs 7 and 10). Then, a mask image (binary edge image) is produced by applying thresholding to the edge image (see Figs 8 and 11). In other words, pixels of the edge image whose value is smaller than threshold  $t_e$  are set to zero and pixels whose value is equal to or larger than the threshold are set to a nonzero value. Figures 8 and 11 show some mask images. Since a video can be viewed as a sequence of frames or fields, the above-stated procedure can be applied to each frame or field of videos. Since the model can be used for field-based videos or frame-based videos, the terminology “image” will be used to indicate a field or frame.

FIGURE 9

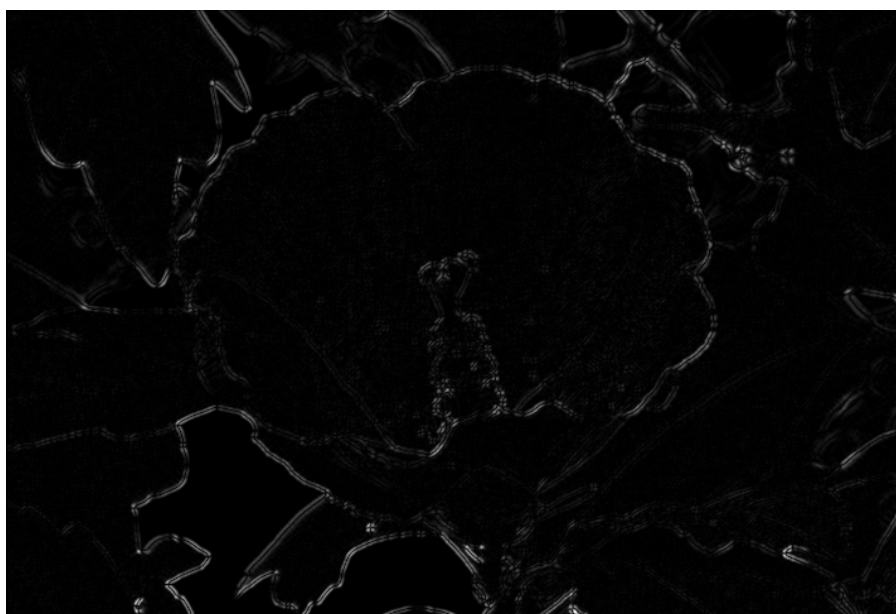
**A vertical gradient image, which is obtained by applying a vertical gradient operator to the source image of Fig. 4**



BT.1885-09

FIGURE 10

**A modified successive gradient image (horizontal and vertical gradient image), which is obtained by applying a horizontal gradient operator to the vertical gradient image of Fig. 9**



BT.1885-10