

RECOMMENDATION 813

METHODS FOR OBJECTIVE PICTURE QUALITY ASSESSMENT IN RELATION TO IMPAIRMENTS FROM DIGITAL CODING OF TELEVISION SIGNALS

(Question 64/11)

(1992)

The CCIR,

considering

- a) that with the increasing application of digital coding and bit-rate reduced transmission, the assessment of coding impairments is of critical importance;
- b) that objective measurements are necessary for specific evaluations as routine supervisions or system optimizations;
- c) that several types of objective measurements may be defined for different digital systems and applications;
- d) that the adoption of standardized methods is of importance in the exchange of information between various laboratories,

recommends

- 1. that the general methods described in Annex 1 for obtaining objective measurements of digital picture quality should be used;
- 2. that the choice of the appropriate objective measurements for a system or an application should be made on the basis of the information given in Annex 1.

ANNEX 1

1. Introduction

With the increasing application of digital coding and bit-rate reduced transmission, the assessment of coding impairments is of critical importance. An understanding of these assessment methods is relevant not only to the performance of new coding equipment, but also to an interpretation of measurements made on such equipments and to specifications for target performance. Moreover digital codecs as with all adaptive or non-linear digital processes, cannot be fully characterized with traditional television test signals or patterns.

Quality of codecs for distribution application can be measured objectively, quality specifications being expressed against the subjective judgement of observers.

Studies indicate the desirability of establishing relationships between objective measurements of signals impaired by digital coding, and the visual quality of the picture thus obtained. This Annex gives progress towards this end, which is proving more difficult to achieve as codec complexity increases.

The quality of a codec designed for contribution applications however, could in theory be specified in terms of objective performance parameters because its output is destined not for immediate viewing, but for studio post-processing, storing and/or coding for further transmission. Because of the difficulty of defining this performance for a variety of post-processing operations, the approach preferred has been to specify the performance of a chain of

equipment, including a post-processing function, which is thought to be representative of a practical contribution application. This chain might typically consist of a codec, followed by a studio post-processing function (or another codec in the case of basic contribution quality assessment), followed by yet another codec before the signal reaches the observer. Adoption of this strategy for the specification of codecs for contribution applications means that the measurement procedures given in this Annex can also be used to assess them.

2. Digital codec classification

The function of digital coding is to reduce the bit rate needed to represent a sequence of images while ensuring minimal loss in picture quality. Coding equipment does this, first by removing as much statistical redundancy from the images as possible (i.e. no loss in quality occurs as a result of this conceptual first stage). Then, if more bit-rate reduction is necessary, some distortion has to be introduced into the picture, although one of the objectives of codec design is to hide this distortion by exploiting certain perceptual insensitivities of the human visual system.

It is convenient to divide codecs into two classes, those using fixed word-length coding and those using variable word-length coding (see definitions in § 3.1 and 3.2 respectively). The latter class is more efficient and complex, and includes all recently proposed systems for coding 4:2:2 video to the range 30-45 Mbit/s. The former class is however sufficient to permit 4:2:2 video to be reduced to 140 Mbit/s while still preserving the quality demanded for contribution applications. A further sub-division of these classes is also useful, into intrafield (or spatial) codecs and interframe (including interfield) codecs, which contain frame (or field) stores permitting them to exploit the redundancy which exists between successive picture frames (or fields).

There is emerging a third class of codec which employs variable word-length coding but which is being designed for variable bit-rate networks. These codecs can in principle, preserve a constant decoded image quality subject to the bounds of peak network demand. The quality-testing of such codecs would have to take into account the nature of the network used and the statistics of the data injected by all of its users, and remains to be studied.

3. Objective assessments of codecs in terms of perceived picture impairments

3.1 Fixed word-length codecs

With fixed word-length codecs a fixed number of bits is used to represent a fixed number of source picture samples. For example in fixed word-length PCM or DPCM codecs, a fixed number of bits is allocated to each picture sample, and in fixed word-length transform or vector quantization codecs, a fixed number of bits is allocated to each block of picture samples.

3.1.1 Methods based on the use of synthetic test signals

In these codecs the impairment introduced into each received picture sample of an image is dependent upon the values of those samples in the locality surrounding it, either in the same field (for an intrafield codec) or in the same and previous fields (for an interframe codec). It is therefore possible, using suitably chosen two or three dimensional digital test signals, to artificially provoke the degradations characteristic of digital image coding.

Some of these degradation factors have acquired names such as false contouring, granular noise, blur, blocking impairments, etc., relating to their interpretation by observers. Having provoked these distortions, their magnitudes can be objectively measured and, using experience gained from subjective assessments these measurements could then be related to some quantification of codec quality. Relating the degradation factors to their interpretation by observers may prove difficult in interframe coding systems or systems employing some adaptive processing because they can vary at any moment, with motion or adaptation of the coding algorithm. In a proposed method, the subjective assessment test first uses scales derived from pairs of opposite adjectives (the semantic differential method), and then the results are analysed by principal component analysis to extract the picture quality degradation factors. The classification results can be tested by applying multiple regression analysis which relates the factors to subjective judgements. A list of picture quality degradation factors is presented in Table 1.