(2)

FINAL REPORT OF
A FEASIBILITY STUDY INTO
THE USE OF IMPAIRMENT UNITS AND
THEIR SUMMATION IN
THE EVALUATION OF VIDEO
DISTRIBUTION SYSTEMS

DEPARTMENT OF SUPPLY AND SERVICES

Philip A. Lapp Limited

HEAD OFFICE: SUITE 302, 14A H
Phone (416) 92
OTTAWA OFFICE: SUITE 904, 280
Phone (613) 23

Government of Canada
Department of Communications
DOC CONTRACTOR REPORT DOC-CR-TS-82-0039
DEPARTMENT OF COMMUNICATIONS - OTTAWA - CANADA
TECHNOLOGY AND SYSTEMS RESEARCH AND DEVELOPMENT

FINAL REPORT OF A FEASIBILITY STUDY INTO THE USE OF IMPAIRMENT UNITS AND THEIR SUMMATION IN THE EVALUATION OF VIDEO DISTRIBUTION SYSTEMS

DEPARTMENT OF SUPPLY AND SERVICES
CONTRACT NO: 1ST82-00095

DOC SCIENTIFIC AUTHORITY: METIN B. AKGUN, Ph.D., P.Eng.
CLASSIFICATION: UNCLASSIFIED

CONTRACTOR

November 1982

Work Carried Out By: K.E. Hancock
Dr. J.N. Barry
C.L. Balko

Approved By: ....
K.E. Hancock
Senior Vice President

SUMMARY

Any video distribution system is subject to impairments of the information being distributed. In practice the technical quality of such a distribution system is determined by first specifying acceptable levels for each impairment of interest, and then measuring the impairments of the actual system against such specifications. Weaknesses of this traditional approach include the fact that the technical measurements have no provision to ensure specific correlation with the transmission of a satisfactory video display to the viewer, and the fact that there is little or no correlation between the individual parameters specified. Viewers of course see an image in which all impairments are present to a greater or lesser extent, and the acceptability of the image is determined by the combined effect of all impairments.

In the past some work has been undertaken, particularly in Europe, to overcome these problems by relating all impairments of a video distribution system to a common unit, the impairment unit (IMP) which in turn is directly related to the subjective impact of the impairment on a viewer. Some preliminary work has also been undertaken on the summation of these impairments to provide an overall evaluation of the picture being viewed. In general this work has related to the PAL television transmission system and is not necessarily applicable to the North American NTSC television transmission system.

This feasibility study evaluated the work previously undertaken to ascertain its applicability to North American systems, defined the gaps in this work and provided a work plan to fill these gaps. Based on this work plan an assessment of facilities, equipment and human resources to do the work was produced.

In addition, the possibility of summing values of impairments to provide a single number expressing the quality of a system was analysed.

Key conclusions reached were that:

- The concept of using subjective impairment units related in graphical form to objective measurements to permit repeatable subjective assessments of electronic distribution systems is valid.

- That this concept has been used by a number of major broadcasting systems on a world-wide basis for a number of years. These users include the Canadian Broadcasting Corporation.

- The use of this approach has improved the cost effectiveness of improving overall network performance.

- The work already carried out on a world-wide basis, and due for publication this year by the CCIR, provides a major basis on which to carry out additional work so that the concept can be used to evaluate cable television systems.

- It is very likely that the concept is applicable to any type of electronic distribution system.

- The use of the concept provides major improvements in cost-effectiveness of evaluating cable television systems.

- In addition, a number of impairments typical of cable television systems, difficult to measure objectively, are amenable to impairment unit evaluation.

From these major conclusions, and others given in the body of the report, the following key recommendations are made:

- That the Department of Communications plan future work that would result in the availability of impairment unit curves to the cable television industry.

- That a viewing laboratory be established either by the Department or through agreement with the private sector to carry out this work.

- That the work be carried out on a step by step basis to minimize investment.

- That further work be considered to advance the state of the art in the use of the <u>summation</u> of impairment units.

- That work be undertaken to ascertain the usefulness of the concept of the evaluation with impairment units for other electronic distribution systems, particularly satellite distribution systems.

- That a training tape using impairment unit curves derived for cable television be produced for use within that industry.

* * *

ACKNOWLEDGEMENTS

In the course of the study considerable co-operation and assistance was received from many groups and persons. In particular, the help freely given by the following was much appreciated:

Dr. C.A. Siocos	-	Vice President, Engineering, CBC
Mr. P.W. Kaiser	-	Manager Quality Control, CBC.
Mr. E. Tsang	-	Communications Research Centre, Ottawa
Mr. R. Poirier	-	Vice President, Engineering, CCTA
Mr. D. Marino	-	United Video Ltd., Ottawa
Mr. L. Mudd	-	Development Manager, Rediffusion Engineering, U.K.
Mr. R. Seacombe	-	Technical Manager, Visionhire Cable Ltd., U.K.

* * *

CONTENTS

CONTENTS (Cont'd)

CONTENTS (Cont'd)

FINAL REPORT OF A FEASIBILITY STUDY INTO THE USE OF IMPAIRMENT UNITS AND THEIR SUMMATION IN THE EVALUATION OF VIDEO DISTRIBUTION SYSTEMS

1. PREAMBLE

Any electronic distribution system is subject to impairments of the information being distributed. Typical examples of such impairments are noise, interference and cross talk. In practice, the technical quality of such a distribution system is determined first by specifying acceptable levels for each impairment of interest and then measuring the impairments of the actual system against such specifications.

In a video distribution system the impairments of interest will be greater in number and will include, but not necessarily be restricted to, noise, hum, cross modulation, intermodulation, ghosts, chrominance/ luminance delay inequality, etc.

As an example cable TV systems in Canada are currently evaluated against BP-23 which specifies specific impairment parameters, defines permissible limits and indentifies acceptable objective measurement methods for these impairment parameters.

This traditional approach has, of course, been developed over scores of years to provide repeatable scientific standards and measurements, thus ensuring that any system meets its technical requirements. It does, however, have at least two significant weaknesses.

The first of the weaknesses is that the technical measurements have no provision to ensure specific correlation to the overall objective of the system, which is to provide a satisfactory video display to the viewer.

Stated briefly, there is no direct correlation between objective specifications and subjective acceptance by the viewer. Specification limits for a particular parameter are frequently selected on a technical basis with only superficial reference to the subjective impact on the viewer. Specification units are in technical terms such as dBs, degrees, time or voltage, etc. Only occasionally are the individual parameter limits specifically related to whether the picture provided is acceptable or not, and under what conditions.

The second main weakness relates to the fact that there is little or no correlation between the individual parameters. The engineer tends to think of such things as signal-to-noise, co-channel interference and intermodulation as separate and specific impairments. The viewer of course sees an image in which all impairments are present to a greater or lesser extent and in which the acceptability of the image is determined by the combined effect of all impairments.

Perhaps paradoxically, this can, in the limit, lead to two opposing effects. It is possible for a system to be over specified, in-so-much that it must concurrently meet worse case limits for each of many technically conflicting parameters, while the presence at the same time of a few key impairments just within specification may provide a less than acceptable video image, as the viewer is seeing the combined effect of these impairments.

These two problems are not, of course, new. To date they have defied a practical general solution, although much work has taken place and there is a reasonable body of literature on which to build any future investigations.

The work to date has fallen into two fairly distinct, and complementary areas. The first of these, carried out mainly in Europe, has related to the defining of a common impairment unit, derived by detailed strictly controlled subjective test to a standard methodology, and the correlation of this unit by graphical methods to the quantified limits of all parameters of interest. Some preliminary work has also been carried out on the summation of these impairments, in other words, the subjective quality of television pictures with multiple impairments. This work has been supported by certain standardization work carried out by CCIR. Much, but not all of this work has been in relation to PAL TV systems.

In parallel with the work described above, a considerable body of research has been carried out, mainly in North America, on the subjective effects of TV impairments.

Much of the work carried out to date on impairment units and their summation is scattered, uncorrelated, and therefore not readily available to the Department as a tool in its consideration of effective methods of evaluating video systems.

While there are certainly gaps in the work that has previously been carried out, these gaps were not defined and there were no plans in Canada for either filling these gaps or ascertaining on a world wide basis whether work is being carried out to fill them. There is also a perceived need in Canada for tested and acceptable graphs of impairment units against technical parameters for the NTSC television system.

There is also a need to develop summation methodology for various combinations of these parameters.

2. NEED FOR A FEASIBILITY STUDY

It is generally agreed that there are significant advantages to the concept of using a common impairment unit (IMP) to define all impairment parameters in a cable TV distribution system. It is envisaged that the use of such units could either be complementary to, or instead of, the various parameters, units and limits currently specified in such documents as BP-23.

In addition, the possibility of summing these impairments to give a figure of merit for a system with multiple impairments would be a major step forward in specifying and testing systems to meet the needs of the actual viewer without costly over-specification and with assurance of always providing an acceptable picture.

Despite the work carried out by Allnatt, Flor, Taylor, the CCIR and others, there were no generally known impairment unit curves available to evaluate the NTSC TV distribution system, nor is there definitive work applicable to the NTSC system permitting summation of impairments.

Before Canadian regulatory and other bodies can use the concept of impairment units and their summation there was a need to evaluate the work carried out to date, assess its applicability to Canadian requirements in a number of fields, to estimate any additional work required to permit impairment units and their summation to be used with confidence in Canada, and to specify such additional work. This study is intended to fill that need.

3. SCOPE

The scope of this feasibility study is defined by the contractual statement of work as follows,
TO:

- Carry out a literature search of work carried out throughout the world on the definition of impairment units and the summation of impairments. (6)

- Analyze this work and determine the extent of the applicability of the results for the evaluation of video distribution networks using the NTSC TV system. (7)

- Evaluate the possibility of summing a number of impairment units, with emphasis on the summation of noise with other individual impairments. (7)

- Evaluate the usefulness of impairment units on the performance evalution of wideband cable systems operating to 400 MHz and beyond, and of cable systems using a mixture of digital and analog signal transmission. (7)

- Comment on the cost effectiveness of impairment unit techniques in the performance evaluation of cable systems as compared to the approach taken in BP-23, from the point of view of the Department and of the cable operator. (12) (13)

- Take into account relevant recommendations of the CCIR and CCITT including but not necessarily limited to CCIR recommendation 500 "Method for the Subjective Assessment of the Quality of TV Pictures" and CCITT report number 486-1 "Transmission Performance of Television Circuits Designed for Use in International Connections." (4) (7)

- Use the information gained from the above tasks to recommend a plan of work required to gain any further information required to permit the Department of Communications to use the concept of impairment units and their summation with confidence in future regulations. (9)

- Detail the requirements for carrying out such work. This should identify human resource and equipment requirements, and determine what specific laboratory facilities are required. These requirements as well as the recommended work plan should allow the Department to proceed with the work in well defined stages. It should also enable the Department to consider accomplishing each task with the Department's own resources and facilities with the aid of outside consultation, or through outside contracting of the entire task. (10)

- Deliver a final report containing the above analyses, work plan requirements and recommendations. (12) (13)

Note: The figures in parenthesis indicate the section(s) of this report relating to the given item in the scope.

4. DEFINITIVE DOCUMENTS USED AS A BASIS FOR THIS STUDY

During the planning for this study the necessity for relating the work to acceptable reference documents became apparent. As the findings will be applied primarily to cable television systems and as a standard method of subjective assessment of the quality of television pictures has been agreed to on an international basis, the following standards and recommendations were selected as the definitive documents used as a basis for this study. Copies of these documents are attached as Appendix A.

4.1 Technical Standards and Procedures for Broadcasting Receiving Undertakings (Cable Television). BP-23, Issue 2. January 1, 1982, Part 4 Technical Standards

BP-23 is the document used in Canada as the regulatory standard for all cable television systems. The parameters, and parameter limits used in this study were selected primarily, but not exclusively, from this document.

4.2 Cabled Distribution Systems Primarily Intended for Sound or Television Signals Operating Between 30 MHz and 1 GHz. IEC Publication 728. June, 1982. Chapter III Performance Requirements

The International Electrotechnical Commission has recently released this standard to provide internationally recommended parameters, limits and methods of measurement for cable TV systems. It has been selected as a definitive document in addition to BP-23 to support that document and to provide information on parameters not covered in BP-23.

4.3 Method of Subjective Assessment of Quality of Television Pictures. CCIR Draft Recommendation 500-1 (MOD F) 1982.

This CCIR Recommendation provides methods of laboratory assessment of picture quality; a grading scale; and standard viewing conditions for the subjective assessment of the quality of television pictures. As such it was considered essential to use this document as a basis for the work carried out in this study.

4.4 Subjective Assessment of the Quality of Television Pictures. CCIR Draft Report 405-3 (MOD F) 1982.

This CCIR Report provides an excellent detailed summary of all major results of objective experiments being carried out in different countries using different methods. It also describes in considerable detail methodologies for television subjective testing providing information on the analysis and presentation of this material.

It also provides laws for the derivation of impairment units from various types of test results giving recommendations as to the methodology to be used in various cases.

In addition this report provides examples of methods of processing the results of subjective tests and details of test procedures.

Finally the document gives information on ascertaining the subjective quality of alphanumeric and graphic pictures.

As such, this document, together with that identified in 4.3 above, provides basic information on setting up and carrying out subjective measurements suitable for the evaluation of all types of impairment, permitting the calculation of impairment units to be evaluated.

5. VIDEO SYSTEM PARAMETERS CONSIDERED

As the initial step in the study, the most relevant impairments of video transmission, distribution and receiving systems were identified, and, in conjunction with the Scientific Authority, a selection was made of those which would be considered during the research and analysis phase of this study.

The selected parameters were broken down into the seven separate groups given below, each group having the same approximate subjective effect on the viewer. For example, those impairments giving a snowy effect were grouped together, as were those giving horizontal bars.

In each group of impairments, a key impairment is identified as being that of greatest concern in video distribution systems.

"SNOWY" IMPAIRMENTS

IMPAIRMENT	EFFECT	IMPAIRMENT PARAMETER	BP 23 UNITS	BP23 LIMITS	KEY IMPAIRMENT OF GROUP
Noise	Snow	Carrier to noise ratio	dB	not less than 40	X
Intermodulation (composite beats)	Grainy snow	Carrier to composite beat ratio	dB	not less than 53	

"HORIZONTAL BAR" IMPAIRMENTS

Hum	Slowly rolling wide horizontal bars	Carrier to hum ratio	dB	not less than 34	X
Direct pickup (Direct pickup signal different program to cable TV transmission. Non-phase locked)	Rolling horizontal bars (venetian blind)	Signal quality	Impairment grade	4*	
Co-channel interference	Rolling horizontal bars (venetian blind)	Signal quality	Impairment grade	4*	
Electrical interference (impulse noise)	Broken horizontal lines	Carrier to noise ratio	dB	not less than 40	

"BACKGROUND PICTURE" IMPAIRMENTS

Crossmodulation	Noise and different program in background	Crossmodulation ratio	dB	not less than 48	
Direct pickup (direct pickup signal different program to cable TV transmission. Phase locked)	Different program in background	Signal quality	Impairment grade	4*	X

"PATTERN" IMPAIRMENTS

Intermodulation (small number of beats)	Moving herring bone pattern on picture	Carrier to beat ratio	dB	not less than 30 (W curve)	X

"GHOST" IMPAIRMENTS

Echos	Right hand edge image (ghost)	Echo rating	%	not less than 7	X
Multipath reception	Right hand edge image (ghost)	Echo rating	%	not less than 7	
Direct pickup (off air transmission same frequency as cable TV	Left hand edge image (ghost)	Echo rating	%	not less than 7	
Chroma delay	Colour only leading or trailing edge image	Chrominance - luminance delay inequalities	Nanoseconds	Plus or minus 150	

"COLOUR QUALITY" IMPAIRMENTS

IMPAIRMENT	EFFECT	IMPAIRMENT PARAMETER	BP 23 UNITS	BP 23 LIMITS	KEY IMPAIRMENT OF GROUP
Differential gain	Colour washout or emphasis	Differential gain of modulator	dB	2	X
Differential phase	Improper colour	Differential phase of modulator	Degrees	Plus or minus 5	
Low picture definition and improper colour	Soft picture with colour bloom or washout	Frequency response of processors (headend)	dB	Plus or minus 1.5	
Low picture definition and improper colour	Soft picture with colour bloom or washout	Frequency response of modulators (headend)	dB	Plus or minus 1.5	

DIGITAL SIGNAL CAUSED IMPAIRMENT

Field blanking interval data interference	Flare disturbance: retrace line: audio buzz	Not specified	-	-	X

* For Class One TV signals only - other signals not specified.

6. LITERATURE SEARCH

Several concurrent approaches to the literature search were undertaken to ensure that all relevant information on the use of impairment units and their summation was obtained. The first approach was to review all literature on the subject currently in the resource files of PHILIP A. LAPP LIMITED. From this significant base many references were obtained and were used in a manual search of the major technical libraries in Ottawa, including the Canada Institute for Scientific and Technical Information. In parallel with this work, two computer bibliographic searches were initiated, the first being an INSPEC search of all literature in that data base dated between the 1st of January 1977 and September 1982 relating to the key words of "television impairment - television interference - television subjective testing". The second computer search used the INSB data base for the period January 1969 to December 1976 using the same key word selection. These computer searches resulted in a total of 442 abstracts being available for review.

These abstracts together with the several hundred references obtained by the manual search were the subject of an initial review which resulted in a selection of a 122 documents which were reviewed in detail.

This detailed review identified much useful information, but perhaps of greater importance, led, as part of the second stage of the literature search, to discussions with key researchers in the field.

Perhaps a paramount of these was an interview with Dr. C. Siocos, Vice President, Engineering, of the Canadian Broadcasting Corporation. Not only has Dr. Siocos considerable experience in the development of impairment units, but he has carried out a great deal work in this field for the CCIR. As part of these activities Dr. Siocos was able to provide copies of CCIR documents not due to be published for several months. Perhaps the most important of these was Report AR-11 which contains impairment unit graphs for a number of impairments for all major television systems. This document and others obtained from Dr. Siocos will be discussed in detail below.

The second stage of the literature search also lead to English translations of certain key documents by Dr. Flor of the Austrian Broadcasting Authority being made available. These documents, together with those reviewed in the first part of the literature search, total 130 documents, and are listed in Appendix B, under five separate headings, each in chronological order. The key documents are identified with an asterisk.

7. ANALYSIS OF THE RESULTS OF THE LITERATURE SEARCH

The first part of the analysis was carried out to determine the extent of the applicability of this world wide work on impairment units and their summation to the evaluation of video distribution networks using the NTSC Television System. This analysis was carried out strictly in relation to those impairments agreed upon with the Scientific Authority as relevant to this work, and which are identified in Section 5 above. This work was carried out in three parts as follows:

- the concept of using impairment units when only one parameter is involved;
- the concept of summing a number of impairment units;
- the concept of using impairment units for the evaluation of wide band cable television systems operating to 400 MHz and beyond, and using a mixture of digital and analog signal transmission.

During the analysis of the results of the literature search, two invaluable CCIR documents came to light. These were CCIR Recommendation 500-1 described in 4.3 above and CCIR Report 405-3 described in 4.4 above. Together these documents form the basis for the standards for a viewing room, the grading scales required in subjective testing and details of appropriate methods of subjective measurement for various circumstances. As such the standards and methodologies given in these documents are recommended as part of all of the work identified in this report as being required to be done in the future, unless it is specifically mentioned otherwise.

7.1 Impairment Units for Single Impairments

It was found that very considerable work has been done on the development of subjective correlation with objective measurements for various parameters pertaining to TV transmission systems. This work resulted in the development of a number of quality scales ranging from 2 point scales to 7 point scales. These finally evolved into the five grade scale accepted by the CCIR in recommendation 500-1 (MOD F). These impairment scales led to the development by Allnatt and others of a method of processing the results of subjective tests to arrive at a unified method of presenting what is called an impairment characteristic, the value of the objective measurement of impairment being related to the mean score of the picture assessment and the "IMP" unit scale. Impairment units and the CCIR five grade scale (as used in BP-23) are directly relatable as shown in Appendix A, Part 4, page 15.

A detailed review of the work leading up to this concept, and of the very many results obtained, would be long, repetitious and inappropriate to this feasibility study. Instead, given below, are the major pieces of work which resulted in the development of impairment unit graphs for each of the specific parameters identified in Section 5 of this Report. With these listings is an analysis of their applicability for use in assessing Canadian cable television systems.

Background information, and the work leading up to the acceptance of the concept of impairment units by many researchers, the CBC, the BBC, British Telecomms and by the CCIR, may be obtained by reading the references given in Appendix B of this Report.

The work identified below in relation to the various groups of impairments, has been selected to be the most advanced available. In many cases a number of earlier papers led up to the results given in the paper identified. These earlier papers have only been included when additional relevant information is contained in them. In this, and succeeding sections, the system to be evaluated has been taken to include the television set, as this is an integral part of a viewer's subject assessment. In particular this refers to sub-sections 9.2.2, 9.3.2 and 9.5.2 and related paragraphs in Section 7.

If it is required to assess a cable TV system as defined in BP-23, i.e. terminating at the subscriber terminal, the TV set, should be preceeded by a converter with an output channel that has been checked as being free of ambient signals.

7.1.1 Impairments Giving a "Snowy" Subjective Effect

In this Section are given abstracts of documents which provide graphs or other information on the use of impairment units for the evaluation of carrier or signal-to-noise ratio, or for the evaluation of intermodulation ratio where the intermodulation contains a sufficiency of multiple high level beats to subjectively appear as "grainy" noise.

- Noise: CCIR Report AR-11 "Subjective Quality of Television Pictures in Relation to the Main Impairments of the Analog Composite Television Signal - Annex - Suggestion of a Possible New Recommendation"(Ref. 4.17). This draft document recommends that the characteristics referred to in it be regarded as a reference for the relationship between picture quality and the objective value of each of the distortions in question, assuming that only one of them is present. It is expected that this recommendation will be accepted at the next relevant CCIR meeting.

  The characteristic in question is for continuous random noise and as such contains only unweighted noise due to uniform spectrum noise (white noise). As such a correction factor is required to make this useful for evaluating carrier to noise for NTSC systems as defined in BP-23. During the literature search two documents (Refs. 5.20 and 5.26) were identified which provide such correction factors.

It is concluded that the work required to confidently use this work is a recalculation of the curve based on references 5.20 and 5.26 to convert the parameter to carrier-to-noise ratio as defined in BP-23 (ref. 5.31), followed by a short subjective test to confirm the conversion.

- Intermodulation (with a large number of high level beats): A study of the published literature did not reveal any work carried out in relation to the use of impairment units for the evaluation of intermodulation containing many beats at a high level giving the effect of "grainy snow". This in itself is not unexpected in-so-much that in the first case most work in this field is related to the broadcast industry where the effect is essentially unknown, and secondly the definition is itself somewhat subjective being a very extreme case of intermodulation interference. The visual effect is however, quite different from that given by the intermodulation of a small number of beats at comparatively low levels.

  The literature has revealed two documents relating to the subjective effects and measurement of intermodulation (References 5.18 and 5.30). It is suggested therefore that original work is required to derive curves for the subjective effects of intermodulation products, both those causing "grainy snow" and those causing a herring bone pattern (see 7.1.4 below). This work will consist of full subjective tests using test set-ups derived from these two references as well as from the CCIR work. Cognizance would be taken of such matters as number of interfering beats, and their levels, giving various intensities of the two types of subjective effects; together with such considerations as synchronization of signals and the use of coherent carriers and frequency allocations.

7.1.2 Impairments Giving a "Horizontal Bar" Subjective Effect.

In this section are considered those impairments which subjectively give some form of horizontal bar, either moving or non-moving, on the television picture. These include impairments caused by 60 cycle hum, which subjectively gives a slowly rolling wide horizontal bar; direct pick up of a non-phase locked program different to that being transmitted over the cable TV system and co-channel interferences, both of which give a "venetian blind" effect. Also included is the large number of impairments which come under the general heading of electrical interference. These include impluse noise, caused by electrical motors and other arcing effects, noise from electrical power lines, radiated interference from other electrical equipment and so on. The perceived effects of these impairments will vary, but as a group they can be described as giving the subjective effect of broken horizontal lines, either moving or non-moving, on the picture.

These impairments are not included in the work covered by CCIR Report AR11. (Ref. 4.17) However, a limited amount of work by various authors on these impairments has been carried out.

- Hum: The literature review revealed very little work on this impairment. It is referred to in Reference 1.1, but only in relation to black and white PAL pictures, and only as the initial development of the work which eventually led to the impairment unit concept. In Reference 3.13 and 3.15 methods of measuring hum modulation and of simulating hum modulation for subjective measurements are outlined. These references however do not relate this work to the concept of impairment units.

  From the above it is recommended that to permit curves of impairment units against carrier-to-hum ratio for NTSC systems to be derived, that a full subjective measurement program be carried out, probably using the measurement method given in Reference 3.13, together with those given by CCIR.

- Direct pick-up of a non-phase locked program different to that carried on the cable TV distribution system: The detailed literature analysis did not reveal any work carried out on this impairment. This is probably due to the fact that it tends to be unique to cable television distribution systems. It is therefore recommended that a full scale subjective test program be carried out to derive curves for this impairment. It should be noted that currently there is no specific technical requirement for this impairment in BP-23 other than that the overall signal shall meet, Impairment Grade 4, for signal-to-noise. It is suggested that the work should include derivation of objective limits of wanted to unwanted signal level.

- Co-Channel Interference: To a large extent the comments relating to direct pick-up also apply to co-channel interference, the subjective effect of the two impairments being essentially similar. With co-channel interference however, a simulation methodology has been derived in Reference 3.13 and summarized in Reference 3.15. Reference 1.16 does provide some early impairment unit work carried out with PAL transmission co-channel interference that should be reviewed prior to new work being carried out. It is recommended that a full subjective test be carried out, probably in conjunction with that for direct pick-up, to provide co-channel interference curves. Again BP-23 does not provide objective limits for this parameter and the comments made under direct pick-up also apply here.

- Electrical Interference (impulse noise): Some initial work on PAL Systems relating low frequency random noise to impairment units was carried out in Reference 1.14. It should be noted that to the extent that this is valid for NTSC Systems, portions of the work also apply to carrier-to-hum ratio.

  As no further work on this impairment has been identified, it is recommended that initially Reference 1.14 be reviewed to ascertain just how much of this work can, with confidence, be related to NTSC Systems, and a subjective test program of an appropriate scope be set up to obtain the information required to provide NTSC impairment to impulse noise curves. It should be noted that BP-23 does not provide a separate objective limit for this type of noise, although subjectively it is usually quite identifable from thermal noise.

7.1.3 Impairment Units Giving a "Background Picture" Subjective Effect

Under this heading are the impairments of cross-modulation and direct pick up of a phase locked signal different from that being carried on the same frequency on the cable television network. Once more neither of these impairments are covered by the current CCIR work, relating as they do solely to transmission systems carrying more than one video channel. However, considerable work has been done on cross-modulation, and its video equivalent, cross-talk.

- Cross-modulation: Reference 1.28 is the main piece of documentation located on true cross-modulation. In addition this work relates to cross-modulation in CATV (cable television) Systems, and uses the North American definition of this parameter. Curves of mean opinion against cross-modulation level were derived from subjective testing, permitting comparatively ready calculation of cross-modulation level in dBs against impairment units. This work, however, relates mainly to the impairment caused by the line synchronization pulses of the unwanted signal, rather than a background picture caused by phase locked or synchronized cross- modulation. The line synchronization pulses cause vertical rather than horizontal bars and this form of the impairment should be considered in any work to be carried out.

  The remaining reference relating to cross-modulation, reference 1.35, whilst giving a impairment unit to cross-modulation ratio curve, states that this curve is derived from grading tests on undistorted cross-talk. It then makes the unsupported statement that the subjective effects on television pictures are similar and therefore the given relationship is applicable. The literature search has not revealed the results of any test to support this statement. References 3.13 and 3.15 provide measurement methods and test set-ups for NTSC cross-modulation subjective tests but do not provide correlation between these tests and impairment units.

  References 1.2 and 1.18 relate to video cross-talk with the former summarizing early work on black and white pictures which helped lead to the concept of impairment units, whilst the latter was later work on PAL colour television signals, and derives mean score (and impairment unit) figures against colour bar cross-talk signals.

In the light of the above it is recommended that the test set-up used in reference 3.13 be used in a CCIR subjective test set up to check the correlation between the figures derived in reference 1.18 and given in reference 1.35, and subjective results for an NTSC signal. Should a minimum series of tests confirm the correlation, further tests would be unnecessary.

- Direct pick up of a phase locked program different to that being carried out on the cable television system: The literature search did not reveal any work relating to this impairment. It is recommended therefore that the confirmation tests recommended for cross-modulation above be repeated using a direct pick-up signal injected into the receiver input, via a splitting device and variable attenuater, to check if there is correlation between the subjective impact of cross-modulation and direct pick-up of this form.

7.1.4 Impairments Giving a Pattern Subjective Effect

This section is limited to one impairment, that of intermodulation with a small number of beats, giving rise to a herring bone pattern of an intensity and colour depending upon the closeness of the beats to the critical carriers of the video signal, as well as the amplitude of these beat frequencies.

The only reference revealed by the literature search (reference 1.29) uses the relationship between the six point impairment scale and impairment units with the protection ratio requirements for the PAL system to derive a curve. The literature search did not reveal evidence of any subjective tests carried out to support these curves. References 3.13 and 3.15 provide simulation set-ups for the subjective evaluation of cross modulation with one or more beats.

In view of the lack of work carried out on this impairment, and its key importance in cable TV distribution systems, it is strongly recommended that a full series of subjective tests be carried out to derive curves of impairment units against intermodulation ratio. These tests should take cognizance of the visual impact of intermodulation of one, and a small number of, interfering beats at various levels and at various distances from critical subcarriers within the NTSC video signal.

7.1.5 Impairments Giving the Subjective Effect of a Left Hand or Right Hand Edge Image (GHOSTS)

Four impairments are included in this group namely; echoes or reflections within the distribution system, which give the subjective effect of a right hand edge image; multipath reception, which gives an identical subjective effect; a direct pick up of an off air transmission of the same program as that being transmitted over the cable television system, which gives the subjective effect of a left hand edge image; and finally chroma delay which gives the subjective effect of a colour-only leading edge or trailing edge colour image.

- Echoes: The CCIR in reference 4.17 provides definitive curves of the impairment characteristic for an undistorted positive echo having a delay of one microsecond together with correction factors to be subtracted from the value of signal/echo ratio for other values of delay. As such, no further work is recommended for the use of these impairment unit curves with NTSC video systems.

- Multipath Reception: The subjective effect of multipath reception is identical to that caused by echoes, particularly long delay echoes within the distribution system itself. Indeed in certain of the definitive work on this impairment, for example reference 1.9, echoes and multipath reception are treated as a single impairment. As such it is recommended that the CCIR curve for echoes described above, be used for multipath reception, and that no further work be undertaken regarding this impairment.

- Direct pickup of an off air program identical to that being transmitted by the video distribution system: The literature search did not reveal any work relating to this impairment. It is typical of video distribution systems linked to inexpensive domestic receivers where the precaution of changing transmission channel of a local signal is not implemented. It applies to a number of small cable television and MATV systems, and is therefore of interest. Though it may be intuitively felt that right hand edge images would have the same subjective impact as left hand edge images, discussions with cable television licencees indicate that this is not necessarily so. It is recommended therefore that a small scale subjective test, using a delay line to provide delays of the wanted signal, be carried out to check the correlation or not of the subjective impact of these two impairments.

- Chroma delay: This impairment is dealt with by the CCIR curves in Report AR/11 (Ref. 4.17), with curves for 525 line and 625 line systems being given. As such no further work is recommended on this impairment.

7.1.6 Impairments Giving a "Colour Quality" Subjective Effect

Under this group of impairments are included differential gain, differential phase and low picture definition and improper colour caused by poor frequency response of headend processors and headend modulators.

- Differential gain: The CCIR Report AR/11 includes a curve of differential gain against impairment units with curves for NTSC, PAL and SECAM systems. As such this is recommended for use in Canada with no further work required.

- Differential phase: Again CCIR Report AR/11 includes impairment unit curves for differential phase for the three major television systems. As such it is recommended that the curve for the NTSC system is used and no further work carried out.

- Low picture definition and improper colour: This impairment is caused by poor frequency response of either the headend processor or the headend modulator. As the subjective effect is identical, both causes will be treated as the one impairment. These impairments are manifestations of linear chrominance-luminance gain inequalities, and the use of the CCIR chrominance-luminance gain inequality curve is recommended, taking into account the somewhat different limits required by BP-23. If this precaution is taken no further work is recommended for these impairments.

7.1.7 Picture Impairments Caused by the Carriage of Digital Signals

Under this heading are those impairments to a video picture caused by the carriage of digital signals in the field blanking interval. The current CCIR recommendations or reports do not cover this impairment.

Although some work has been done in this field, particularly with regard to the impairment caused by digital codecs (References 1.37 and 1.38), considerable work remains to be done. The work relating to codecs came to the conclusion that the impairments arising from the use of such codecs are very small and give rise to problems of subjective differentiation of such small changes in impairment. Refined methods of subjective tests have been proposed for this work, using the so called double stimulus method (Reference 1.38 and 4.13). However, little or no work has been carried out regarding other mechanisms of digital impairment of a video signal. While it is understood that certain work in this field has been carried out by the IEC by its radio communications receiver sub-committee (IEC/SC12A), it was not possible to obtain copies of this work during the duration of this study. It is recommended therefore that work in this area be incorporated as part of any future work on the use of impairment units in Canada.

7.2 The Summation of Impairment Units

As mentioned in the preamble to this report the viewer sees an image in which all impairments contributed by the system are present to a greater or lesser extent, and in which the acceptability of the image is determined by the combined effect of all impairments.

It can therefore be concluded that if the subjective effect of these combined impairments can be assessed in a repeatable manner, it would be of very great advantage to the design, construction and evaluation of video distribution systems. Impairments caused by various mechanisms within the distribution system can be allocated in a manner that, whilst giving an acceptable picture to the viewer, will permit the most cost-effective design and construction of a system. In addition the task of evaluating a system will be made much easier and less expensive. Subjective assessments by trained technicians could be evaluated on a comparative basis with far more confidence than is possible today without the use of such a subjective evaluation tool.

It is not surprising therefore that over the last 17 years considerable work has been carried out in an attempt to attain effective and definitive methods of summing impairments.

In reviewing the work carried out in this field a number of conclusions can be drawn:

- Although considerable definitive work has been carried out on the summation of impairments, (see Section 2 of Appendix B) there is not yet agreement between researchers as to the most effective methodology, in particular as to the most effective mathematical distribution model for the subjective opinion range that will provide best correlation in summation of impairments.

- Much of the work carried out in summation is related to the summation of pairs of impairments.

- There is some evidence (Reference 2.4) that a "normal" distribution might be preferred for NTSC systems.

- Professor Flor in Reference 2.8 gives some evidence that a normal distribution is perhaps the best approximation for single impairments whilst that proposed by Allnatt in Reference 1.24 might be more applicable to the summation of impairments. It should be noted however that Flor's work related solely to PAL Systems.

- In discussions with Dr. Siocos, Vice President, Engineering, CBC, it was ascertained that, for practical purposes, the CBC finds the summation of impairments very useful in allocating trade-offs in system design. However, the approach should be used with judgement, and is most useful when small impairment numbers are summed, rather than when the summation of gross impairments is attempted.

In modern cable television systems it is perhaps true to say that the limiting factor is the major increase in intermodulation products as large numbers of TV channels are carried concurrently. This effect becomes apparent after approximately twelve TV channels are distributed, and increases in an exponential manner.

In view of the above, it is recommended that some initial trials be carried out to evaluate the various summation models identified above using two key impairments; wideband RF noise, and intermodulation with a small number of beats. It is suggested that such work, which whilst limited in itself, could provide sufficient information on which to base a detailed plan of work to permit the concept of summation of impairments to be used with confidence in the evaluation of broadband video systems.

7.3 The Use of Impairment Units for the Evaluation of Cable Systems Operating to 400MHz

The literature research did not reveal any documentation relating to the use of impairment units for cable television systems operating to 400 MHz and beyond.

However the work carried out during this study, together with knowledge and experience of modern cable systems, permits some general comments on this concept.

Intrinsically the concept of the use of impairment units is valid for any type of electronic distribution system. In particular it appears to have the potential of being a very valuable tool in the assessment of broadband distribution systems with large numbers of viewers, where it is difficult or impractical to carry out objective tests with sophisticated test equipment on all or many of the terminals, or to close down networks whilst measurements are being taken.

Broadband systems operating to, and beyond, 400 MHz will carry very large numbers of video channels concurrently, and will be very prone to intermodulation impairments giving the subjective effect of superimposed "herring bone patterns", and in severe cases grainy noise. These impairments are particularly amenable to the use of impairment units to evaluate their subjective impact. In addition there is some evidence, but not yet decisive evidence, that the summation of pairs of units such as noise and intermodulation is practical and able to relate back to objective measurements.

One of the most powerful techniques currently being used to permit extension of broadband video systems to a bandwidth of 400 MHz and beyond is that of the selection of coherent carriers. This technique, in which the carrier frequencies of all video channels carried are harmonically related, reduces the visual impact of the beat frequencies caused by the intermodulation mechanism. One result of this technique is, however, to change the subjective effect of the intermodulation impairment to that of one or more background pictures. Little or no work has been carried out on the subjective impact of this type of impairment, particularly where more than one background picture is present. It is recommended therefore that future work take cognizance of this lack of information.

7.4 Concluding Remarks

In this and in earlier chapters we have reviewed and analyzed the work to develop impairment units over the past three decades, work undertaken in major laboratories such as the British Post Office, Bell Telephone Laboratories (BTL), the Canadian Broadcasting Corporation, and others in countries such as the United States, U.K. and Canada. In addition to work done by common carriers in these countries a significant amount of work has been reported by the cable industry to aid the regulatory bodies in their respective countries in monitoring and controlling cable systems that operate under national standards. The result of the effort in these various countries has served not only the particular interests of each country or private organization that performed it, but has been the basis for the work of the CCIR and CCITT to establish internationally acceptable standards by which interoperability of national systems can be assured or at least examined and compared. As has already been demonstrated, this coordination of information in the CCIR has led to concrete results in the development of acceptable impairment units to be used in international systems assuming that it is to be used internationally for national systems. In section 9 of this report, the basic details of a plan of work DOC could undertake to develop or modify impairment units for cable television system monitoring and control are given.

8. A SPECIAL CASE OF VISUAL IMPAIRMENT

In modern cable television systems a significant proportion of alpha numeric programming is carried. Some major cable television systems may have three or four channels devoted solely to such programming, providing broadcast news, TV program information, community notices, weather information and the like.

This type of programming is extremely vulnerable to most types of impairment, in-so-much that there is a clear immobile background, and little movement of the alpha numeric information to provide screening of impairments. In particular it is vulnerable to impairments causing echos or smearing and to colour effects impairing the alpha numeric information itself. Such impairments, as well as being caused by the effects identified in Section 5 of this report, are frequently caused by the equipment generating the alpha numeric programming, and by inappropriate filtering in processors and modulators.

As far as is known there has been no specific consideration of this type of impairment either by the Department, or by other researchers. At this time it is merely identified as a subject for further consideration, particularly as this type of programming is likely to proliferate in the future.

9. WORK PLAN

The analysis of impairment parameters in the previous chapter included a brief suggestion or recommendation for each parameter of the additional work required to permit the use of impairment unit curves in the evaluation of

Canadian video distribution systems, in particular cable TV systems. In this section these recommendations are elaborated upon in sufficient detail to form the basis of a future plan of work. For this discussion, the major breakdown of parameters used in the previous chapters has been retained. At the end of the section a schedule against which experiments and tests might be conducted, is given, bearing in mind the complexity and cost of certain tests.

9.1 Impairments Giving a "Snowy" Subjective Effect

9.1.1 Noise

RECOMMENDATION: Use the work referred to in References 5.20 and 5.26 to recalculate the noise weighting curves to make them compatible with the definitions in BP-23. Implement this filter and follow-up with a short subjective test to confirm the conversion.

PLAN OF WORK: Coordination of the various noise curves in References 5.20 and 5.26 is readily accomplished. At the present time, the cable industry and DOC use as a BP-23 noise reference a flat noise spectrum at RF, 4MHz wide. Since this is a pre-detection measurement, the method yields a carrier-to-noise figure, a measurement much easier to accomplish in the field. Hence its use in BP-23. Common noise references, pre-and post-detection are summarized in Reference 5.20 as follows:

$$C/N \text{ (NCTA) (\& BP-23)} = \frac{\text{(RMS power of RF carrier during synch. pulse)}}{\text{(RMS noise power in 4 MHz wide channel)}}$$

$$C/N \text{ (TASO)} = \frac{\text{(RMS power of RF carrier during synch. pulse)}}{\text{(RMS noise power in 6 MHz wide channel)}}$$

$$S/N \text{ (EIA)} = \frac{\text{(Voltage difference synch. tip to reference white)}}{\text{(RMS noise voltage, 10 kHz to 4 MHz, EIA weighting)}}$$

$$S/N \text{ (CCIR)} = \frac{\text{(Voltage difference blanking pulse to reference white)}}{\text{(RMS noise voltage 10 kHz to 4 MHz, CCIR weighting)}}$$

$$S/N \text{ (BTL)} = \frac{\text{(Voltage difference synch. tip to reference white)}}{\text{(RMS noise voltage, 10 kHz to 4 MHz, CCIR weighting)}}$$

The EIA, BTL and CCIR noise weighting curves take account of the more objectionable features of low frequency noise in a television picture, hence the objective S/N measurement is more closely related to the subjective picture quality.

In Reference 5.20, the relationship between these definitions has been calculated by the author, T.M. Straus. Straus reports tests confirming his predicted relationship between NCTA pre-detection carrier-to-noise and CCIR post-detection weighted signal-to-noise to within 1 db for vestigial sideband modulation:

$$S/N\ (CCIR) = C/N\ (NCTA) - 0.2\ db.$$

Straus also tested his TASO versus NCTA calculations experimentally with good agreement. In Reference 5.20. Clayton derives the relationship between pre- and post-detection noise measurements for FM modulation systems and relates them to the CCIR, BTL and EIA curves.

To implement this plan of work including the subjective test would require minimum time assuming facilities were available, since no unknown quantities are involved. Essentially, it would be an experiment to confirm Straus's work relating NCTA C/N and CCIR S/N. A 'live' test using the configuration of Figure 1 would probably suffice. Consultation with representatives of the cable industry in Canada would be necessary to ensure that the industry would support and use such an impairment unit for noise interference.

9.1.2 Intermodulation (with a large number of high-level beats)

RECOMMENDATION: Undertake original work to derive curves for the subjective effects of intermodulation products, both those causing grainy snow and those causing a herring bone pattern. (See Section 7.1.4 above)

WORK PLAN:* The work of Hood in Reference 5.30 indicates that composite triple beat due to intermodulation, that is a large number of high-level beats caused by third order intermodulation, manifests itself as a distribution of beats in the vicinity of a picture carrier. The beats are so closely spaced that they give the appearance of noise with a shape approximating a normal distribution.

---

*The work plans for the two effects - grainy snow and herring-bone - are discussed separately.

The distribution is not wide; Hood notes that the majority of composite triple beat spectral patterns observed experimentally are approximately 30 kHz wide. Hence the grainy snow appearance on a TV screen.

A systematic attack on the problem would begin with an examination of the true shape of the spectral curve of the beat energy with measurements of the randomness of intensities with time. Close examination would have to be made of the actual nature of the signals inside the composite triple beat envelope to decide whether in fact they are best represented by noise or by a mixture of coherent and stochastic signals. Other input parameters to be taken account of are the number of beat products making up the interference spectrum, their relative magnitude and the dependence of the overall spectrum shape on carrier frequency allocations.

The production of triple beats for subjective tests could be accomplished using a short cascade of overloaded amplifiers, or, alternatively, the effect could be simulated using a filter with a particular type of noisy signal passed through it. The test set-up(s) are illustrated in Figure 2. The work described here might be expected to span several months because of the need for discussion and interaction with the cable industry to make sure that the effects were representative of what is likely to be found in cable systems.

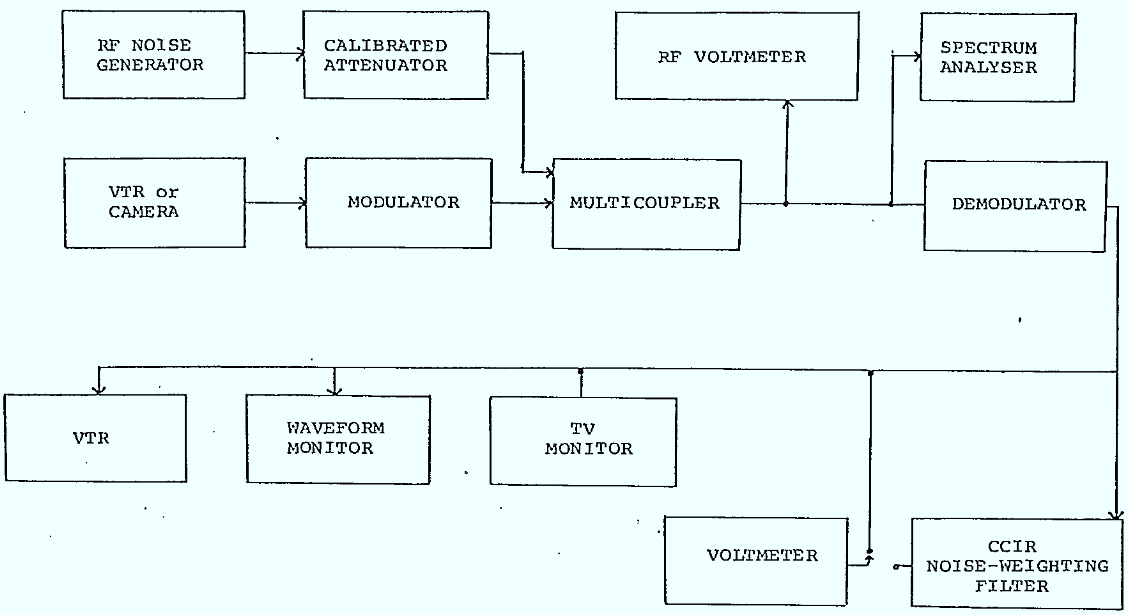

FIGURE 1: VALIDATION OF CCIR NOISE IMPAIRMENT RATING WITH NTCA C/N MEASUREMENT

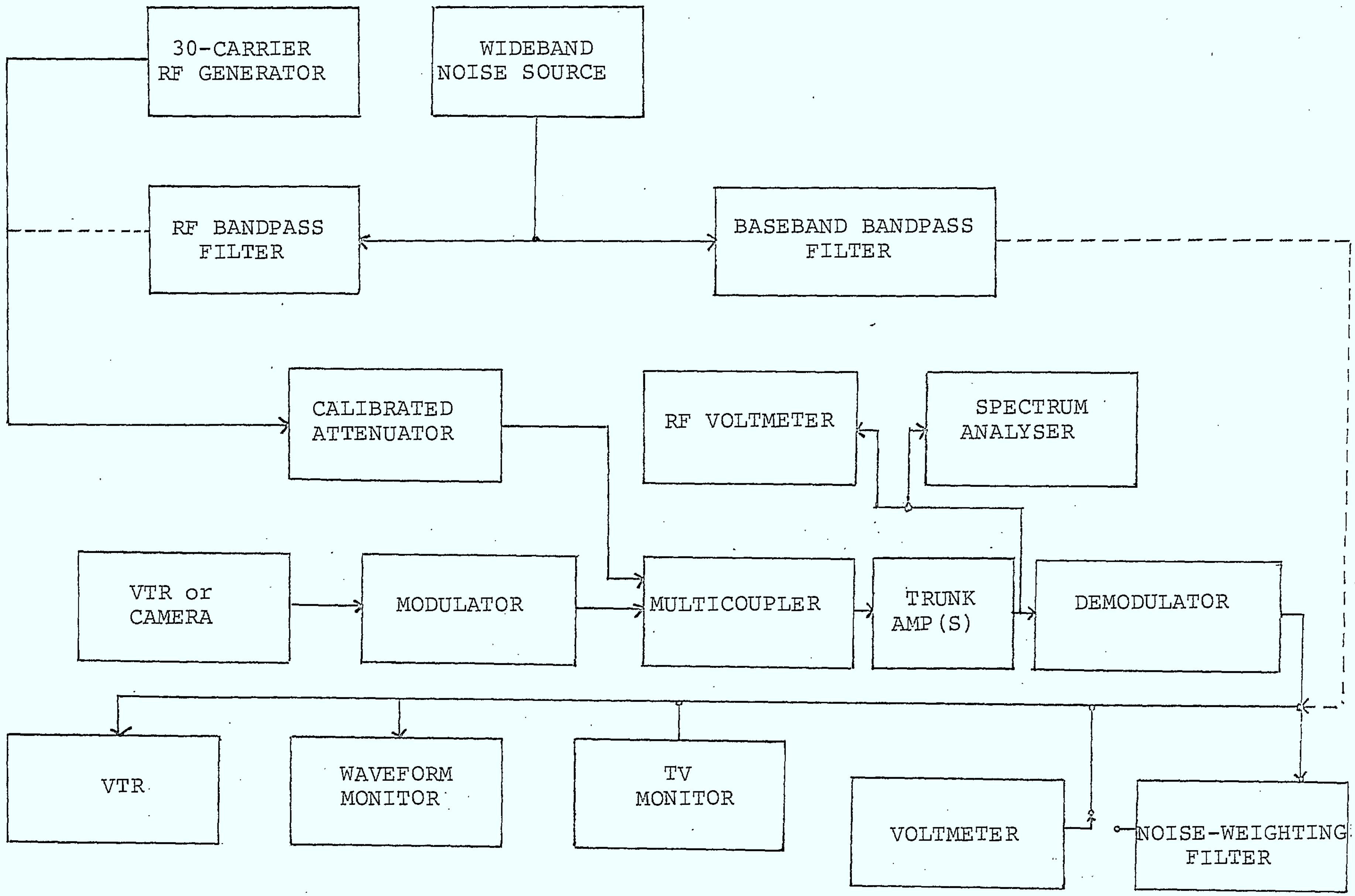

FIGURE 2: COMPOSITE TRIPLE BEAT SIMULATION

With close coordination with the cable industry an "illustrative" impairment unit applicable in carefully defined situations could be defined. This limited impairment would then be an inducement to other workers to build upon this work so as to widen its applicability or alternatively to use it as a supplementary criterion for measuring the effect of this particular type of distortion in Cable Television systems with a large number of channels. Currently, BP-23 requires 53 db suppression of composite triple beats spaced less than 1 MHz from a picture carrier. From subjective tests this is the point at which composite beats are just discernible. No information exists on how the subjective impairment varies with frequency-offset or intensity.

9.2 Impairments Giving a "Horizontal Bar" Subjective Effect

9.2.1 Hum

RECOMMENDATION: Carry out a full subjective measurement program probably using the measurement method given in Reference 3.13 together with those given by CCIR.

PLAN OF WORK: In operating Cable Television systems, and in BP-23, hum control and hum monitoring is accomplished by measuring the amount or percentage of hum as amplitude modulation on an RF carrier. To be consistent with the accepted Cable Television practice, hum simulation impairment should be studied with the hum introduced at RF.

The estimated difficulty of developing an impairment unit for hum is minor. Modulation of the AGC circuit of an RF amplifier would be a satisfactory method of simulating hum impairment. The 60 Hz hum source should have an incremental frequency control so that the hum interference on the television picture could be made to move vertically a slow rate. The impairment unit could be developed over a very short period of time because the physical phenomenon which causes it is well defined and varies little from one Cable Television system to another. Note that a subjective assessment leading to an impairment unit is based on a post-detection signal. Therefore it may be possible to relate pre- and post-detection hum parameters accurately enough to permit repeating the subjective tests with direct hum cross-modulation of a composite video signal.

9.2.2 Direct Pick-Up on a Television receiver of a Non-phase-locked Program Different to that Carried on the Cable TV Distribution System

RECOMMENDATION: Original work is required to derive objective limits of the relative strength of wanted and unwanted signals at an agreed point within the television receiver, for example at the output of the IF demodulator.

PLAN OF WORK: See 9.2.3 below.

9.2.3 Co-Channel Interference

RECOMMENDATION: Conduct full subjective tests, probably in conjunction with that for direct pick-up, to provide co-channel interference curves.

PLAN OF WORK:* Direct pick-up and co-channel interference are readily simulated using a locally generated RF channel and adding an off-air channel at the 'same' frequency to it in the tuner of a television receiver or a demodulator. It is normal practice in coordination of frequency assignments for television stations to off-set channels by several kHz to avoid the more strident effects of co-channel interference which is carrier beats in the order of a few hundred Hz. Hence any test aimed at representing a real life situation would have the carrier separated by some 10 to 25 kHz. As the strength of the interfering carrier is increased the interfering phenomenon is very characteristic and becomes quite pronounced as the interference level increases. The equipment needed to conduct the experiment is relatively straightforward and is illustrated in Figure 3. The main problem in establishing either an objective or a subjective unit lies in relating the subjective impairment of the wanted signal to the frequency separation of the two carriers, their frequency stability, and their amplitude. The effect of the interference to the observer is more annoying and more pronounced as the difference frequency between the two carriers falls below a few kHz. Furthermore, two carriers that are well separated in frequency will show less variation in the interference effect (variation is in itself a source of annoyance) as their individual frequencies either jitter or drift. Careful coordination with the objectives of the Cable Television industry would be needed to establish acceptable subjective and

---

* Plan of work for direct pick-up of non-phase-locked channel is same as for co-channel.

objective measurement situations that would produce results useful to DOC and the Cable Television industry. It will be appreciated that while the objective measurements are quite staightforward over a wide range of different frequencies and levels, the subjective effect is highly dependent on the difference frequency as noted above. For this reason it might be necessary to develop an impairment unit which had a weighting function related to the frequency offset of the interfering carrier. Such a curve has already been suggested in CCIR Report AR11 cited earlier for echo impairments. Alternatively, the effects of frequency-offset and amplitude might be regarded as a summable pair of impairments.

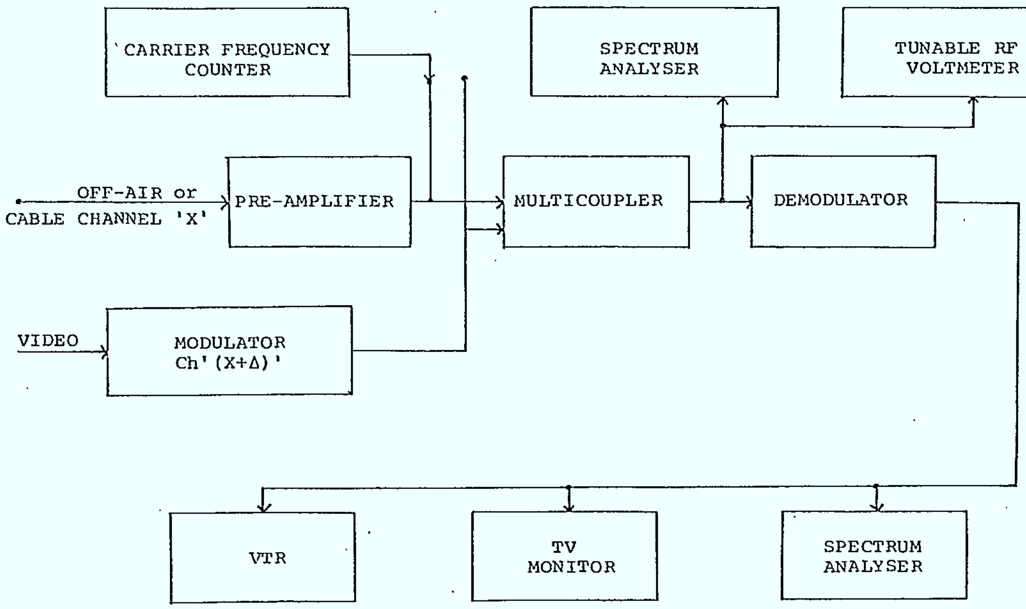

FIGURE 3: SIMULATING IMPAIRMENTS DUE TO DIRECT OFF-AIR PICKUP AT A TV RECEIVER OR CO-CHANNEL INTERFERENCE AT A CABLE HEADEND

9.2.4 Electrical Interference (Impulse Noise)

RECOMMENDATION: Review the work of Allnatt and Bragg in Reference 1.14 to determine how much of this work could with confidence be related to NTSC systems. Devise a subjective test program with sufficient scope to correlate the work of Allnatt and Bragg to provide impulse noise curves suitable for NTSC systems.

WORK PLAN: The article by Allnatt and Bragg treats noise arising from DC to DC converters used in transistor amplifiers for video transmission on co-axial cables in the UK. In the UK systems it proved impossible to completely screen the converter signals from the video path, so a specification for maximum interference level became important. It can be related to impulse noise at RF carrier frequencies insofar as repetition frequency is concerned.

Common sources of impulse noise in North American systems are Hydro power line corona discharges, which cause broadband noise across the entire RF spectrum; ignition noise from internal combustion engines without spark suppression in the ignition system; and noise from electrical motors.

The main problem is to devise a set of subjective experiments to evaluate the effects of impulse noise in general given the variability of the repetition frequency, even though interference from high voltage power lines tends to be at power line frequencies and so has some of the characteristics of hum. As a plan of work it is suggested that a study be made to catalogue published statistics on the spectral and temporal characteristics of this type of interference. There is a growing body of work available on the spectral qualities of arc discharges on satellites to add to the literature on power-line and ignition impulse noise. After classification and cataloguing the interference as to duration, frequency content, etc., it should be possible to identify a predominant characteristic which could then be the basis of an impairment parameter measurement. This result could be compared to Allnatt's impairment measurements using narrow-band noise at 3.2 and 7.5 kHz and sine-wave interference at 1 and 3 kHz and 1 and 3 MHz. Realization of a test set-up for subjective measurements would be relatively straightforward once a consensus was reached on the predominant spectral and temporal features of the interfering signal. A typical simulation set-up is illustrated in Figure 4.

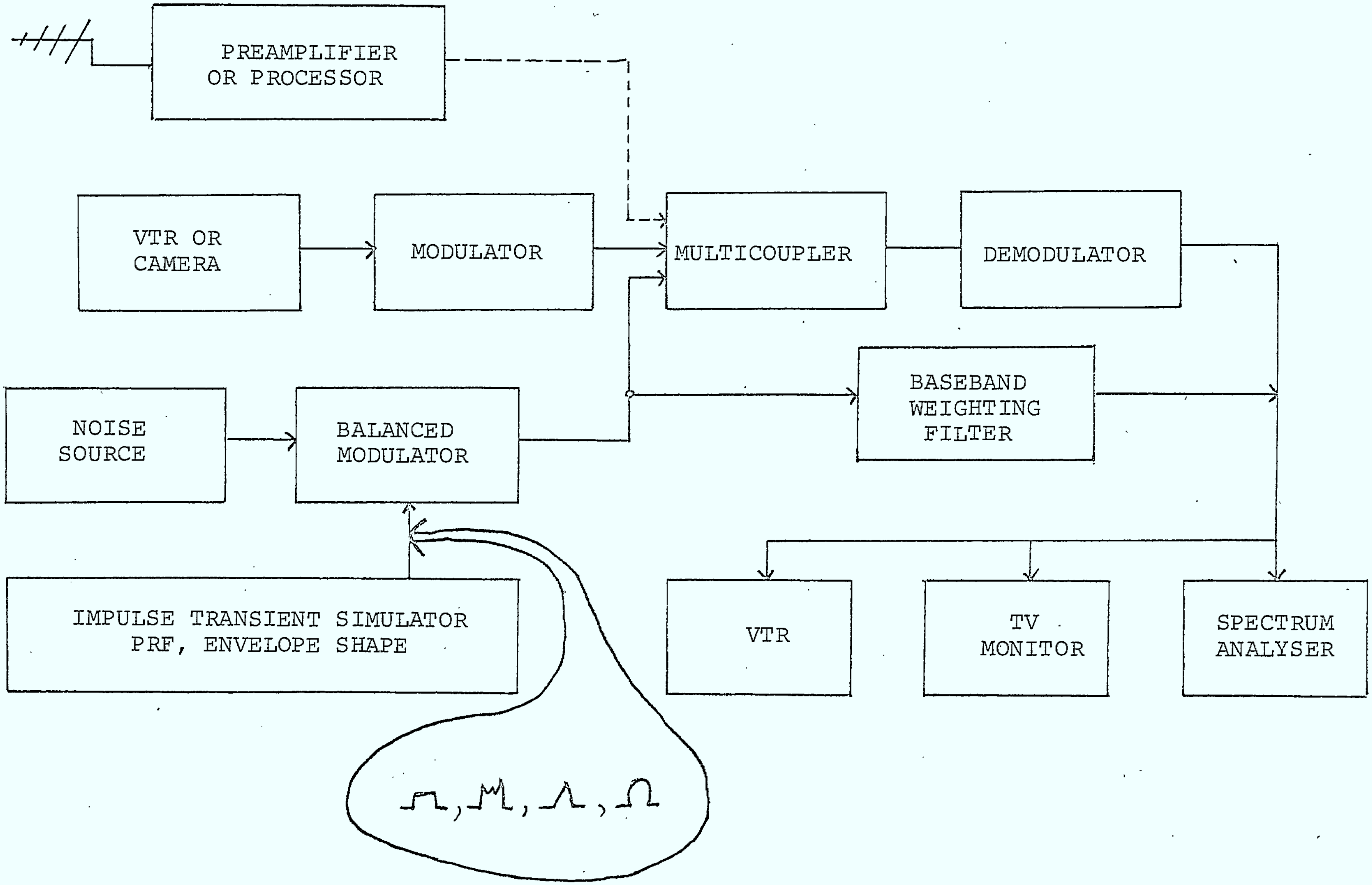

FIGURE 4: SIMULATION OF IMPULSE NOISE IMPAIRMENT

9.3 Impairments Giving a "Background Picture" Subjective Effect

9.3.1 Cross-modulation

RECOMMENDATION: Conduct a test using Mambo's set-up in Reference 3.13 to simulate the effect. Subjective tests should be as recommended by CCIR Recommendation 500. Check the correlation between the figures derived in the work of Corbett, Taylor and Allnatt in Reference 1.18 and the work reported by Street in Reference 1.35. Use an NTSC signal for the test. If the results of Corbett et al agree with Street no further work would be necessary.

WORK PLAN: It is noted that Corbett's work applies to true cross modulation on an NTSC signal whereby the use of negative modulation produces cross modulation in the form of a vertical bar due to the high RF amplitude of the interfering channel during the horizontal blanking pulse channel. Cross modulation is a non-linear distortion so the effect of the blanking pulse is overemphasized with respect to the rest of the video signal. Street's work is more properly described as video crosstalk because it is a linear impairment. In the crosstalk situation the normal balance between blanking pulse and normal video of the interfering signal is preserved so that the vertical bar is not emphasized over the intensity of the interfering picture that goes with it. The subjective effect of crosstalk can therefore intuitively be expected to be noticeably different from cross modulation, making the verification of Street's unsupported statement mandatory.

The procedures for simulating cross modulation as described by Mambo in Reference 1.13 is very straightforward. The RF carrier in an amplifier is simply modulated negatively and non-linearly by a video signal from another source, the objective being to have the interfering video signal cause the wanted carrier to be suppressed slightly during the peaks of the interfering signal. A simple horizontal synch. pulse might even suffice. In a test set-up the line frequency rate of the interfering signal should be adjustable over a small range so that the vertical bar could be made to move slowly across the television screen. Frequency control of horizontal scanning frequencies in North American television broadcast systems is currently so precise that uncoordinated horizontal scanning frequencies in different systems can be expected to differ by no more than one or two cycles in several minutes. Typically the vertical bar in a picture due to cross modulation from an interfering channel, with no synchronization of horizontal frequencies, will require up to 30 minutes to move from one side of television screen to the other. To compare and evaluate the work of Corbett et al with that of Street, it would be necessary to relate the subjective impairment of the vertical bar due to cross modulation, where there was only a faint interfering scene, to the same vertical-bar impairment when accompanied by an interfering scene of proper relative strength. Slow sideways motion of the vertical bar should be provided for in both tests.

9.3.2 Direct Pick up of a Phase-Locked Program Different to that being Carried out on the Cable Television Network

RECOMMENDATION: Repeat the confirmation tests recommended for cross modulation using a direct pick up signal injected into the input of the monitor. Check the correlation between the subjective impact of cross modulation and direct pick up of this form.

PLAN OF WORK: This linear type of interference is unique to Cable Television systems because two signals are operating on exactly the same frequency, differing only in the picture content and, perhaps, a very small amount in the horizontal line synchronizing rate. The subjective effect is identical to the video crosstalk discussed above in 9.3.1. Confirming Street's results vis-a-vis those of Corbett will produce the same impairment-versus-objective measurement relationship.

9.4 Impairments Giving a Pattern Subjective Effect

9.4.1 Intermodulation with a Small Number of Beats

RECOMMENDATION: Conduct a full series of subjective tests to derive curves of impairment units against intermodulation ratios. The tests should take cognizance of the visual impact of intermodulation between critical carriers and sub-carriers within the NTSC video signal format producing one or a small number of interfering beats of various levels and separated by various frequencies.

PLAN OF WORK: The work required to accomplish the above recommendation is extensive. It bears a relationship to the work plan described earlier for measuring the impairment due to co-channel interference, at least from the point of view of the intensity of the impairment. In both instances the presence of a strong monochromatic signal, in one case the result of beat products, and in the other case the result of an unwanted carrier, produces or can produce a similar effect. Allnatt's work with sine-wave interference in baseband (see 9.2.4) is also relevant. The same problems must therefore be solved; i.e. to achieve a consensus on the frequency of the interfering signals to be used to simulate the effect to the observer. Here again a frequency-offset weighting curve may be needed to apply an impairment standard. The added complication of having more than one interfering frequency increases the complexity of the task in the intermodulation products case. Test procedures to simulate the impairment for subjective evaluation are relatively straightforward. They require a small number of RF oscillators with or without modulation at the horizontal line rate, tunable through at least one TV channel. The difficult part of the measurement would be classification of the observers' reactions to the impairment as the interfering frequency and amplitude changes. The simulated impairments would involve a very large number of scenes due to the mixture of different impairing frequencies and different impairing levels and combinations of them with the complexity of more than one interfering frequency at a time being a contributing factor. Therefore an important initial contribution to the eventual development of an impairment unit would be an indication of how well-correlated individual assessments were, using a large number of admixtures of numbers of interfering signals, amplitude variations, and frequency offsets. The test set-up is in Figure 5.

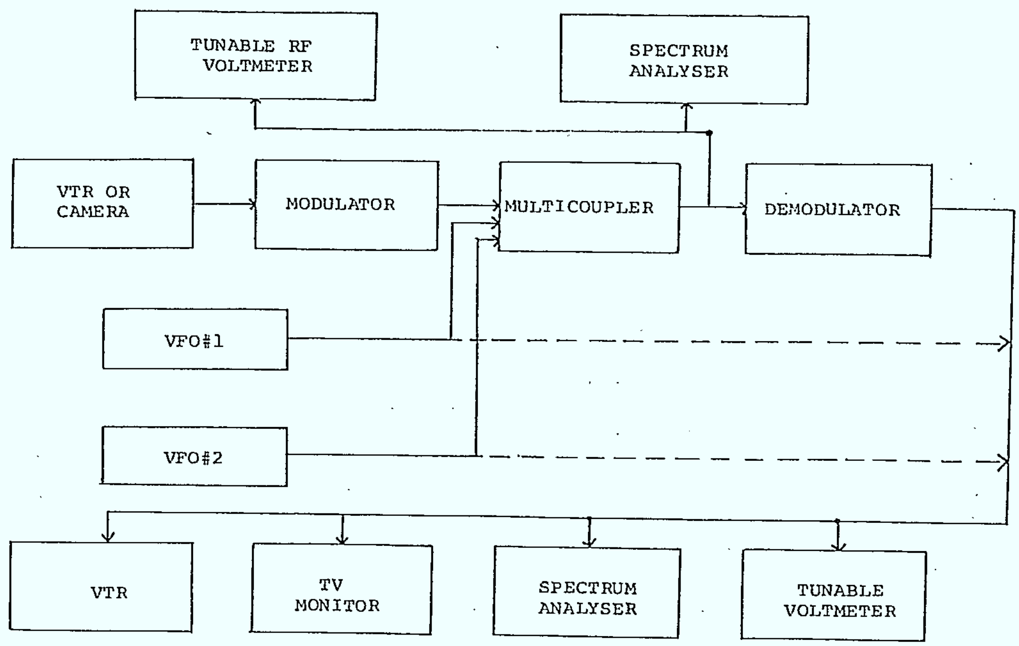

FIGURE 5: SIMULATION OF SECOND-ORDER INTERMODULATION BEATS

9.5 Impairments Giving the Subjective Effect of a Left Hand or a Right Hand Edge Image (Ghosts)

9.5.1 Echoes

RECOMMENDATION: No further work.

PLAN OF WORK: N/A.

9.5.2 Multipath Reception

RECOMMENDATION: No further work.

PLAN OF WORK: N/A.

9.5.3 Direct Pick Up of the same Program Off Air

RECOMMENDATION: Conduct a small scale subjective test using a delay line to provide delays of the wanted signal. Check the correlation between subjective impairments due to leading and trailing echoes.

PLAN OF WORK: This measurement is very straight forward and is readily accomplished with studio instrumentation. The optimum plan of work would see duplication of some of the work used to develop the impairment unit for trailing echoes, followed by work using the same observers but using leading echoes. Although the absolute impairment levels for each of the impairments might not be the same for the same delay and relative intensity, the same weighting curve would be expected to apply. If not, at least a reliable weighting curve can be developed for leading echoes as was done in the case of trailing echoes. This would complete the work on this impairment.

9.5.4 Chroma Delay

RECOMMENDATION: No further work. Use CCIR impairment curves as described in Ref. 4.17.

PLAN OF WORK: N/A.

9.6 Impairments Giving a 'Colour Quality' Subjective Effect

9.6.1 Differential Gain

RECOMMENDATION: No further work. Use CCIR curves in Ref. 4.17.

PLAN OF WORK: N/A.

9.6.2 Differential Phase

RECOMMENDATION: No further work. Use CCIR curves in Ref. 4.17.

PLAN OF WORK: N/A.

9.6.3 Low Picture Definition and Improper Colour

RECOMMENDATION: No further work. Use CCIR curves for chrominance-luminance linear inequalities as given in Ref. 4.17.

PLAN OF WORK: N/A.

9.7 Picture Impairments Caused by the Carriage of Digital Signals

RECOMMENDATION: Include work in this area as part of any future work in the use of impairment units in Canada.

PLAN OF WORK: Measurements of impairments due to digital VBI signals, to be useful, must relate to those digital signals that can, if uncontrolled, give rise to significant visual or audio impairment. The control being exercised currently by broadcasters using these signals is such that any picture impairment attributable to their use is virtually imperceptible. Until such time as these signals become more complex, in their spectral or temporal make-up, extensive research into rating impairments in a cable television system is not warranted. Accordingly, accessing recent IEC work plus maintaining a watching brief on the increasing use of VBI signals is most appropriate at the present time. When the impairment due to more complex signals is perceptible enough to reduce the picture quality to, say 4, on the CCIR 5-grade quality scale, meaningful measurements could be undertaken. Note that complex signals in the VBI could conceivably affect the quality of the audio signal in a cable television channel, but the magnitude of the effect cannot be estimated until detailed characteristics of the complex VBI digital signal causing the impairment are known.

9.8 Impairments in Alpha Numeric Scenes

This special case has been pointed out in Section 8. No specific work is recommended at this time because most of the perceptible impairments can be traced to the quality of alpha numeric character generators at the headend of a cable television system. The quality of future presentations, will improve as the quality of the equipment available at reasonable cost to a cable operator improves. It is noted here that television broadcasters can normally provide unimpaired alpha numeric scenes because they can justify the use of expensive equipment to maintain constant quality in the program into which the alpha numerics are inserted. Note also that an alpha numeric display, or a full-field static test signal, provided it is definitely known to be impairment-free at the point of origin, is a powerful diagnostic tool for assessing the performance of a television transmission channel. However, the extraordinary visibility of impairments in such a static or quasi-static scene presents other difficulties, in relating what is observed to the subjective performance of the channel for the carriage of signals for which it is intended.

9.9 Progression of Work

Based on the work carried out in Sections 8 and 9 above, a summary of the current status of impairment unit curves relating to NTSC systems for all of the impairments specified in Section 5 was drawn up and is given in Figure 6. However a detailed plan of work must take into account priorities and the order in which various tasks are to be undertaken. In order of importance, bearing in mind that cost-effective improvements to the methods of evaluating cable systems is the main issue, noise, and intermodulation, especially high-order intermodulation products are the two impairments over which a cable operation has sole control, the two impairments that figure heavily in system design trade-offs, currently and in the future and, consequently the two impairments most likely to yield a long-term benefit.

Co-channel interference would rank third in importance not because it is not a severe impairment, but because it usually affects only a relatively small number of channels in a Cable Television system.

FIGURE 6

WORK REQUIRED FOR FULL SET OF CABLE IMPAIRMENT CURVES

REF. NO.	IMPAIRMENT	NO WORK	CONFIRMA-TIVE TESTS	FULL SUB-JECTIVE TESTS	COMMENTS
9.1.1	Noise		*		Assessments themselves straightforward and short. Expensive, high-quality components required to reveal subtle changes.
9.1.2	Intermodulation (with a large number of high-level beats)			*	Two phases of work: 1. Research phase: analyse and simulate a generally-acceptable impairing signal. 2. Assessment phase: straightforward.
9.2.1	Hum			*	Assessment is straightforward and short. Introduce hum impairment at RF and at baseband and correlate assessments. Expensive set-up as in 9.1.1.
9.2.2	Direct pick-up on a television receiver of a non-phase-locked program different to that carried on the Cable TV distribution system.			*	9.2.2 and 9.2.3 in conjunction. Assessment is complex due to interaction of amplitude and frequency-offset of impairing signal.
9.2.3	Co-Channel Interference			*	
9.2.4	Electrical Interference (Impulse Noise)			*	Two phases of work, as in 9.1.2 above.
9.3.1	Cross-modulation		*		Assessment is straightforward and brief. High quality set-up is required.
9.3.2	Direct pick-up of a phase-locked program different from that being carried on the Cable TV network.		*		Similar to cross-modulation (and crosstalk) in 9.3.1.
9.4.1	Intermodulation (with a small number of beats)			*	Assessment is complex. Same comments as in 9.2.2 and 9.2.3.
9.5.1	Echoes	*			
9.5.2	Multipath Reception	*			
9.5.3	Direct pick-up of the same program off-air.		*		Assessment is straightforward and brief. High quality set-up is required.
9.5.4	Chroma delay	*			
9.6.1	Differential gain	*			
9.6.2	Differential phase	*			
9.6.3	Low picture definition and improper colour	*			
9.7.1	Digital signals in the VBI			See Comments Column	Full tests required if and when impairments in operating systems become perceptible.
9.8.1	Impairments in alpha numeric presentations	*			Charcter generator is itself a source of impairment.

Hum impairment is rare in a well maintained Cable Television system.

Echo in Cable Television systems is being progressively eliminated due to the greatly improved characteristics, and low thermal variation, of critical hardware.

Multipath to a head-end can usually be eliminated by the proper synthesis of antenna lobes.

Direct pick-up of a signal by a home TV receiver can frequently be traced to improper use of local antennas in conjunction with the cable service, or, even more frequently, improper splicing and coupling of several components to the cable signal. With modern TV receivers impairments of all kinds due to direct pick-up are therefore more a matter of educating Cable Television subscribers on how to connect the cable to a TV receiver properly than a matter of measuring the impairment due to an improper (or illegal) connection.

A second but no less important consideration in a plan of work is the orderly acquisition of test equipment. Impairments that can be measured to the most subtle degree in otherwise excellent quality TV pictures can only be portrayed using expensive broadcast quality equipment. During the course of this study, viewing of a training tape prepared by CBC to illustrate impairments required a very high quality video tape recorder and colour TV monitor. Inferior equipment would itself have caused at least one whole impairment level on a 5-level scale. Indeed, during the tests with the CBC tape, slight misalignment of the colour TV monitor resulted in none of the test scenes being rated as impairment-free. In a logical series of tests therefore, those tests which can be performed satisfactorily with moderately priced equipment should be undertaken first. Improvements to the quality and complexity of the tests can be introduced as necessary.

As a final comment on the need to identify very subtle impairments, it is noted that Cable Television systems customarily deliver a signal that is free of significant impairments but is not, as a general rule, impairment free. Broadcasters on the other hand strive for impairment-free signals at point of origin and, if the transmission channel is very clear, at the point of delivery.

From the foregoing considerations the following order of priority of work is recommended as best meeting the objectives of DOC and the CATV industry.

A (i) Perform a critical analysis of composite triple-beat spectra with a full range of even - and odd-order non-linearities in a typical system.

(ii) Develop a respresentative 'test' signal that best emulates this type of interference in future systems employing over thirty channels.

(iii) Repeat (i) and (ii) for impulse noise.

(iv) Examine channel frequency assignments that would either (i) reduce composite triple beat spectral density or (ii) cause it to be located predominantly in one part of the VHF spectrum, narrow enough that frequency assignments in that part could be avoided.

(v) Concurrently with (i) and (ii) assemble the necessary test components to subjectively evaluate single and two-frequency beat interference. Relate the combined impairment of frequency-offset and amplitude. Moderately-priced equipment is sufficient for a first assessment.

(vi) Repeat (iv) with simulated co-channel impairment.

(vii) Conduct subjective tests on the composite triple beat using the signals and equipment in (ii) and (iv) above respectively. If possible do the experiments at baseband to simplify the test system.

B. (i) Discuss the work of Straus and Clayton with the CATV industry to seek agreement on the use of subjective noise impairment units to evaluate CATV systems. If agreement is reached plan its introduction to the industry at large through CCTA and proposed modifications to BP-23.

(ii) Upgrade the quality of the test system and repeat Straus' experiments to confirm the impairment C/N relationship. Repeat for impulse noise.

C. Conduct hum impairment tests.

D. Repeat tests A (iv), (v), and (vi) with the sophisticated system.

10. FACILITIES, EQUIPMENT AND HUMAN RESOURCES

10.1 General

The equipment and facilities necessary to implement the plan of work described in the previous chapter are readily available from a number of sources. The configurations suggested for the various tests follow in principle both CCIR recommendations and current Canadian Cable Television practice and regulations. In this section, the equipment requirements are summarized and an overview of a test facility is provided. The section begins with a description of a basic facility required for reliable subjective testing.

10.2 Facilities

Before discussing in detail the facilities required to carry out the work to complete NTSC impairment unit curves for all of the parameters identified in Section 5, it is perhaps useful to discuss the concept of subjective assessments and their relationship to objective measurements in reference to the provision of facilities to carry out the subjective assessments.

The various CCIR documents, and other references, mentioned further on in this section specify the key parameters of ambient lighting, optimum viewing room size, minimum number of observers, competence of observers, and a number of other parameters required to produce reliable subjective assessments. Although not specifically stated, these requirements lead to the conclusion that the needs for repeatable subjective assessments are best met by the establishment of a permanent viewing laboratory consisting of a viewing room meeting CCIR recommendations, equiped with fixed, calibratable basic viewing equipment, together with a equipment room housing the electronic equipment required for the subjective assessments.

Only by the provision of a fixed and standard facility, with ambient conditions and reference picture quality that can be measured and thus repeated, together with a fixed environment for the viewers, can subjective assessments of adequate confidence level be made.

Such a viewing facility need not be extensive or expensive, although of course the quality of the basic viewing equipment must be high. What is important is that the facility maintain its ambient conditions and its basic viewing equipment on a continuous basis, rather than the equipment, or indeed the room, being used and modified for other work between sessions.

The discussion on required facilities, developed in the remainder of this section, is based upon the premise that such fixed viewing facilities would be developed and used.

The first requirement of such a facility is that the environment in which the subjective testing and viewing is done must be free of distractions in the way of noise, lighting, or physical discomfort to the viewers. Key parameters defining the ambient lighting, the settings on the television monitor and method of presentation are given in Table 1. The values for these parameters have been taken from CCIR Recommendation 500. Two tables in addition to CCIR are also given: the conditions for laboratory assessment for subjective readings as practised or proposed in other countries - Table 2; conditions for assessment during program transmission for experts in OIRG and in Canada - Table 3. Tables 2 and 3 have been taken from CCIR Report 405-3 (Mode F) "Subjective Assessment of the Quality of Television Pictures". The conditions stated in Table 1 are recommended since they provide the best compromise among the various techniques and judgement factors used in a number of countries. CCIR recommends that subjective experiments should be carried out with non-expert observers who have been appropriately introduced to the test method, the grading scale and the impairments to be assessed. The minimum number of observers in a test is 10, with 20 preferred. The grading scale to be used is given in Table 4. Where it is desirable to use a comparison scale instead of an absolute scale, the CCIR recommendation is to use the scales shown in Table 5.

About five still pictures and sequences including moving subject matter should be used. In anticipating a set-up for the work recommended it is assumed that duplicates of scenes that have been used successfully in other laboratories such as these of the CBC would be used, or if new scenes were to be created, they would contain the essential features that are judged to be important in those previously used. For example sample picutres chosen to study the effects of noise should have a preponderance of grey in them since it is known that noise is most discernible against a medium grey background. The choice of pictures for noise measurements is otherwise not important.

TABLE 1

CONTROLLED PARAMETERS DURING SUBJECTIVE IMPAIRMENT TESTING

Condition	Field frequency	50 fields/s	60 fields/s
a	Ratio of viewing distance to picture height	6 (1)	4 to 6
b	Peak luminance on the screen (cd/m²)	70 ± 10 (2)	70 ± 10
c	Ratio of luminance of inactive tube screen (beams cut off) to peak luminance	<0·02	<0·02
d	Ratio of the luminance of the screen when displaying only black level in a completely dark room, to that corresponding to peak white	approximately 0·01	
e	Ratio of luminance of background behind picture monitor to peak luminance of picture	approximately 0·1 (3)	approximately 0·15
f	Other room illumination	low (4)	low
g	Chromaticity of background	white (5)	D_{65}
h	Ratio of solid angle subtended by that part of the background which satisfies this specification, to that subtended by the picture	>9	

(1) Normally 6; if a different ratio is used, this should be stated.

(2) Normally (70 ± 10) cd/m², or (220 ± 30) asb (6), but certain tests may require luminances outside the tolerances, for example, because of flicker, defocusing, etc.

(3) If the ratio is greater than 0·1, the chromaticity has to be nearer to Illuminant D_{65} (7).

(4) The specification is loosely phrased here, since the precise value is not critical, provided it does not conflict with condition *c*.

(5) Not very critical. Any white in the region between standardized illuminants A and D_{65}. See, however, Note (3).

(6) 1 apostilb (asb) $= \frac{1}{\pi}$ candela per square metre (cd/m²).

(7) Illuminants standardized by the International Commission on Illumination (CIE); see International Electrotechnical Vocabulary, Group 45, No. 45-15-145.

The pictures and impairments should be presented in random sequence with the proviso that the same picture (i.e. test scene or sequence) should never be presented on two successive occasions with the same or different levels of impairment.

When using the quality or impairment scale, the range of impairments should be chosen, wherever practicable, so that all grades are used by the majority of observers; a grand mean score (averaged over all judgements made in the experiment) close to 3 should be aimed at, to standardize results.

A session should not last more than roughly half an hour, including the explanations and preliminaries; the test sequence could begin with a few pictures indicative of the range of impairments; judgements of these pictures would not be taken into account in the final results.

(Source: CCIR Draft Rec. 500-1 (MOD F))

TABLE 2

CONDITIONS FOR LABORATORY ASSESSMENTS

Reference	UK {CCIR, 1963-66a}	EBU, OIRT {CCIR, 1963-66c and d}			Fed. Rep. of Germany {CCIR, 1963-66b}			Japan {CCIR, 1963-66e and 2966-69a}		USA {CCIR, 1963-66f}	USA {CCIR, 1966-69b}	Italy {CCIR 1966-69c}	People's Republic of China {CCIR 1978-82}
Observers Category Number	Non-expert 20-25	>6			Non-expert >10			Non-expert 20-25		Non-expert approx. 200	Expert >10	Non-expert approx. 20	Non-expert 250 expert 65
Grading Scale Type Number of grades	Quality 5 (Note 1)	Impairment 6 (Note 9)	Quality 6 (Note 7)	Comparison 7 (Note 11)	Impairment 5	Quality 5	Comparison 5	Impairment 5 (Note 3)	Quality 5 (Note 2)	Quality 6 (Note 8)	Impairment 7 (Note 10)	Comparison 5 (Note 5)	Quality 5
Test pictures Number	4-8	5			>5			>3		2-8	3-4	6	5
Viewing conditions Ration of viewing distance to picture height	6	4-6			6			6-8		6-8	4	6	4-10
Peak luminance on the screen (cd/m^2) (2)	50	41-54			50			Approx. 400 (monochrome) 74-84 (colour) (Note 12)		70 (Note 12)	170 (mono -chrome) 34 (colour) (Note 12)	50	85
Contrast range of the picture	Not specified				Not specified			30/1 to 50/1		Not Specified			
Luminance of inactive tube screen (cd/m^2)	<0.5	0.5			<0.5			Approx. 5 (monochrome) 0.7-2 (colour)		2		Approx 0-5	
Luminance of backcloth (cd/m^2) (3)	1 illuminant C				Approx. 2.5 (1)			Not specified					
Room illumination, average (lx)	3							Not specified		6-5			
Presentation	Random sequence of pictures and impairments				Random sequence of pictures and impairments			Random sequence of pictures and impairments		Random sequence of impair -monts	Random sequence of impairments	Random sequence of pictures and impairments	Random sequence of pictures and impairments

(Source: CCIR Draft Rep. 405-3 (MOD F))

TABLE 3

CONDITIONS FOR ASSESSMENT DURING PROGRAMME TRANSMISSION

Reference	OIRT [CCIR, 1966-69d]		Canada [CCIR, 1966-69e]
Observers Category Number	Expert 1 or 2		Expert 1 or 2
Grading scale Type Number of grades	Impairment 6 (Note 9)	Quality 6 (Note 7)	Impairment 5 (Note 4)
Pictures Type	Television programmes		Television programmes
Viewing conditions			
Ratio of viewing distance to picture height	4-6		4-6
Angle of view, from a line normal to the face of the monitor			$<30°$
Luminance, on the screen, at reference white (cd/m²)			70 ± 7
Chromaticity of the screen at reference white			Illuminant *D*
Luminance of the inactive tube screen	Adapted to the ambient illumination		As low as practicable
Luminance of "light surround" (cd/m²)			10·5 ± 3·5 (Note 14)
Chromaticity of "light surround"			Illuminant *D*

(Source: CCIR Draft Rep. 405-3 (MOD F))

TABLE 4

Five-grade scale	
Quality	Impairment
5 Excellent	5 Imperceptible
4 Good	4 Perceptible, but not annoying
3 Fair	3 Slightly annoying
2 Poor	2 Annoying
1 Bad	1 Very annoying

TABLE 5

Comparison Scale	
3	Much better
2	Better
1	Slightly better
0	The same
-1	Slightly worse
-2	Worse
-3	Much worse

In Reference 1.10 Allnatt notes that the pictures displayed were 40 by 30 cm and they were displayed on 52.5 cm (21 inch) monitor. This would imply that there was approximately one half inch of unilluminated border on the picture tube around the picture being shown. It seems necessary to establish this small unilluminated border to meet condition 'c' of Table 1 which specifies the maximum ratio of luminance of the inactive to the peak luminance of the active areas of the screen. As an alternative, reference blanking on the video can be used.

The CCIR makes no reference to the monitors to be used for subjective tests. A number of the reports cite the use of more than one television monitor or in some cases ordinary television receivers to simulate a variety of receivers. For the purposes of standard tests that would eventually be useful to the CCIR it is suggested that one high quality 52.5 cm (21 inch) monitor be used with possibly a second placed along side for comparison.

A small carpeted room with a homelike decor with one horizontal dimension of at least 5 metres, but with no dimension greater than 7 metres, is required for a viewing room. This room would be free of all apparatus except the television monitor which is displayed prominently along or close to one wall on which there is a white cloth back-drop. The monitor could be either mounted in front of the wall or flush mounted in the wall. The room is equipped with indirect fluorescent lighting at a colour temperature of 6500°C, on the wall behind the monitor. Other lighting is low and also set at 6500°C. For television transmissions in which the quality of the sound is also to be studied, the room should be equipped with independent head sets for viewers who can adjust the listening level to their own comfort. All play-back and control instrumentation and equipment is located in an adjacent area. Chairs that can be set for a viewing distance from 4 to 6 times picture height are located in the center of the room. For a 52.5 cm (21 inch) monitor 6 times viewing distance is approximately 180 cm. The ambient noise level in the room should be low but the room should not be so dead acoustically as to be distracting to the viewers.

10.3 Equipment

The basic equipment needed to prepare for and conduct subjective tests is summarized in Figure 7. The summary includes a television camera but does not include the variety of 'props' necessary to prepare test scenes.

It is assumed that, in the beginning, duplicates of scenes used by other workers in the field would be used. Furthermore, the variety of methods by which these scenes would be produced is such that it would be best left to the personnel in charge of the program to select the method best suited to what is available at the time. One assumption has been made as to the preparation of the viewing material. It is assumed that the material being viewed will be pre-programmed in advance and will not be interrupted in any way by anyone in charge of the tests. That is to say, while the person in charge of the tests may be present in the viewing room, this person will not have any control in the sense of interrupting or replaying any of the scenes. Such a flexibility would erode the credibility and uniformity of the tests unacceptably. The implication here is that all testing material will be carefully produced with high-quality editing, voice announcements, and the like, into a single video tape. The only exception would be the early low-quality testing of the continued impairment of frequency-offset and amplitude in single- and double-beat interference. These tests could be done 'live' the first time. In the event that the signal-to-noise ratio or other characteristics of tape recorders are not sufficiently controlled to permit presentation of a "perfect" picture, a live" program" of televised transmission of the test pictures may be necessary in some instances. In such a test, it would be necessary to pre-record all announcements so that the cueing would be completely unobtrusive.

10.3.1 Equipment Detail

In specifying the key pieces of equipment required to carry out the work plan, the approach has been taken to initially discuss in broad terms the requirements for each major piece of equipment. This sub-section concludes with a list of the major equipment required to carry out the tests. This list includes key specification parameters, the quantity of equipment required and the tests in which it will be used.

Colour Television Camera - A studio quality colour television camera is required to prepare video tapes of original, unimpaired scenes. Purchase of this expensive component could be avoided if suitable video tapes could be obtained from, for example, the Canadian Broadcasting Corporation or some other worker in the field. Preparation, under contracts, by a professional production house is another option.

Video Tape Recorder (VTR) - A broadcast quality video tape recorder is required to reproduce tests of the various scene impairments and level of impairment to be

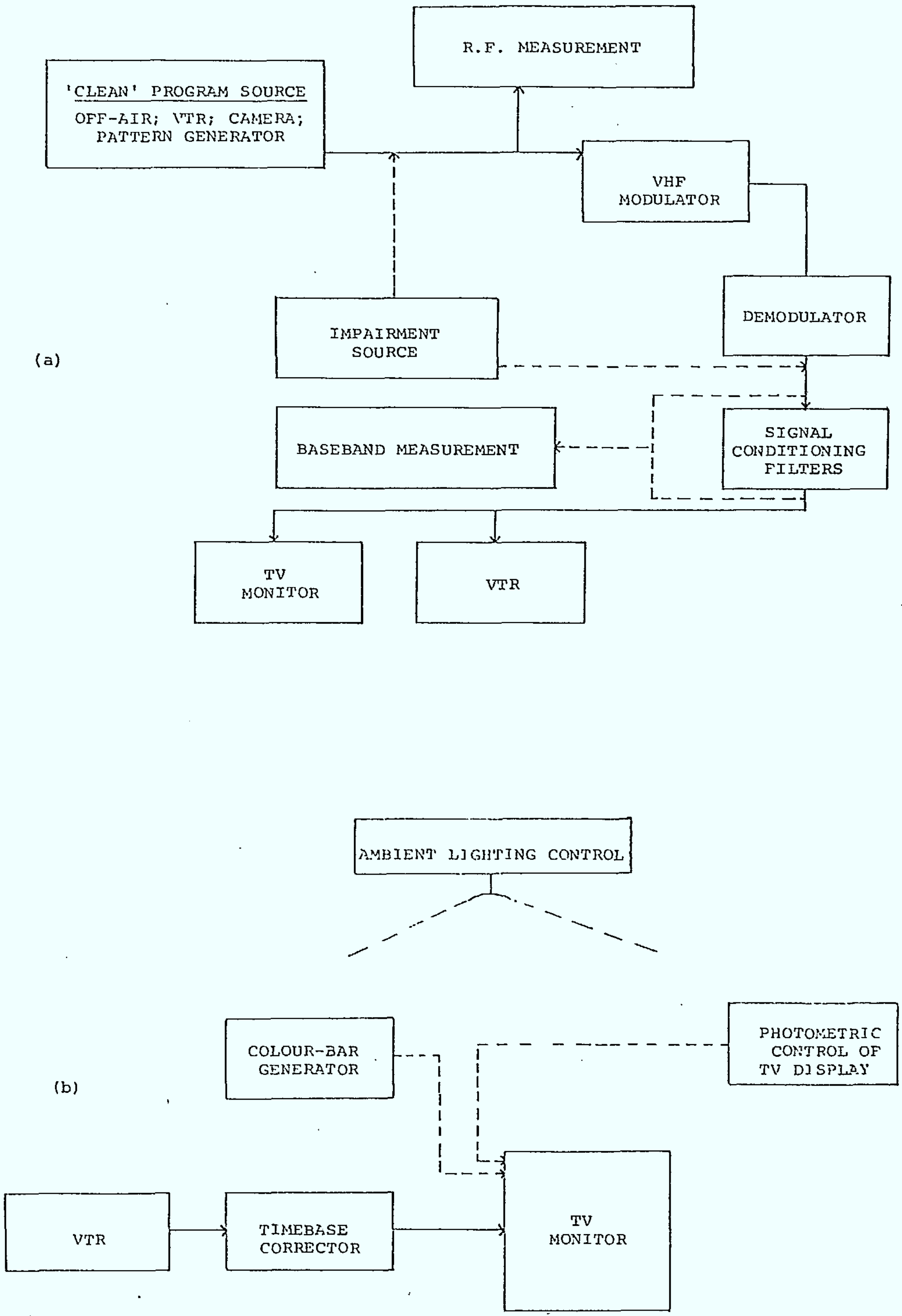

FIGURE 7 : TEST SET-UP FOR STUDYING TELEVISION IMPAIRMENTS
(a) Production (b) Subjective Assessment

studied by the viewers. A time base corrector is necessary to remove distortion due to recording and play back. It may be necessary to acquire a second high quality VTR on a temporary basis so that unimpaired pictures may be presented to an RF modulator into which the impairments are introduced, the RF signal being subsequently demodulated and re-recorded with the impairment present in a second VTR. For this particular set-up a high quality television camera could be substituted for the VTR but it would be more difficult to duplicate exactly the original working material - the scene.

Television VHF Channel Modulator - A modulator is required to permit the linear addition of interfering monochromatic RF signals (simulating beats), shaped noise and hum to VHF channels.

VHF Television Channel Demodulator - A demodulator is required to faithfully demodulate the impaired signals that have been developed in the RF channel.

Noise Source - The noise source is necessary to simulate the effect of noise at RF and baseband. Therefore the noise source should cover the spectrum from a few hundred hertz to at least 100 Mhz.

Weighting Filters - For particular tests low pass filters designed to meet the CCIR and EIA noise weighting characteristics will be required.

Colour Television Monitor - A studio quality monitor is required to view impaired and unimpaired scenes.

Television Colour Analyser - A precision colour analyser is required to assure day-to-day "objective setting" of colour monitors.

Spectrum Analyser - A high quality spectrum analyser with wide dynamic range is required to measure levels of interference introduced into a modulated VHF or IF channel.

Test Signal Generator - A signal generator is required as a source of waveforms and synchronizing pulses for colour bars and other VITS full-field signals needed to set up the monitor and, in general, align the system.

Editor - A simple editing console is required to fade and insert voice and picture signals during tape production.

Wave Form Monitor and Vectorscope - A wave form monitor and vectorscope are required to examine and measure certain impairments as they appear at the output of the demodulator.

VHF/UHF Antenna - A high quality VHF or UHF antenna is required to receive a local off-air signal.

TV Signal Processor/Preamplifier - A TV signal processor or a preamplifier is required to raise the level of the off-air signal and, if necessary, convert it to another channel for subsequent use in a test set-up.

Baseband Distribution Amplifiers, RF Amplifiers - Amplifiers are needed to restore signal levels after attenuators and multicouplers in test loops. An RF amplifier is used in an overloaded condition to produce intermodulation.

Frequency Selective Volt Meter - A frequency selective volt meter is desirable to confirm and refine the measurements made on the spectrum analyser.

RF Signal Source - A multiple-carrier RF generator will be needed to simulate the interference due to interfering RF signals in a television channel. The source should be capable being amplitude modulated at the horizontal synchronization rate.

Test Oscilloscope - A test oscilloscope is necessary for general set-up and service of the various combinations of instruments used for the tests.

Miscellaneous - A number of passive switching networks, attenuators, resistive mixers, and multicouplers will be needed to prepare the test material and monitor tests in progress.

MAJOR EQUIPMENT LIST

EQUIPMENT	QTY	USE IN REF #	KEY SPECIFICATION PARAMETERS
Colour Television Camera	1	All	NTSC; Studio quality
Video Tape Recorder	1	All	NTSC; Studio quality
Time Base Corrector	1	All	NTSC
VHF Channel Modulator	1	All	
VHF Channel Demodulator	1	All	
Broadband Noise Source	1	9.1.1, 9.1.2, 9.2.4	1 kHz to 100 MHz
Video Distribution Amp.	2 (min.)	All	DC-6 MHz; 2 I/P; 4 O/P
Wideband RF Amplifier	2 (min.)	All	30-100 MHz
Colour Television Monitor	1	All	NTSC; studio quality
Television Colour Analyser	1	All	CCIR

MAJOR EQUIPMENT LIST (Cont'd)

EQUIPMENT	QTY	USE IN REF #	KEY SPECIFICATION PARAMETERS
Video Test Signal Generator	1	All	NTSC
Waveform Monitor	1 (min.)	All	NTSC
Vectorscope	1 (min.)	All	NTSC
Spectrum Analyser	1	All	20 Hz - 10 MHz
Spectrum Analyser	1	All	50 - 350 MHz
Video Production Editor	1	All	3 video ports, 1 audio port
VHF Antenna	1	9.2.3, 9.3.1	Single VHF channel
VHF Signal Processor	1	9.2.3, 9.3.1	Single VHF channel
VHF Preamplifier	1	9.2.3, 9.3.1	Single channel
Multiple-Carrier RF Generator	1	9.1.2, 9.2.3, 9.4.1	Minimum 36-channel simulataneous, with H-rate AM
Low Pass Filter	2	All	0 - 4.2 MHz; 0 - 6 MHz
Low Pass Filter	1	9.1.1, 9.1.2, 9.2.4	CCIR noise weighting
Low Pass Filter	1	9.1.1, 9.1.2, 9.2.4	EIA noise weighting
RMS Voltmeter	1	All	0 - 6 MHz
Selective RF Voltmeter	1	All	30 - 350 MHz
Oscilloscope	1	All	General Purpose; DC-10 MHz; Dual channel
Tunable Band Pass Filter	1	All	60 Hz - 4 MHz
Multicoupler (Resistive Adders)	2	All	30 - 350 MHz
Calibrated Attenuator	2 (min.)	All	DC - 350 MHz
Terminating Switch	2	All	DC - 10 MHz; 4 I/P; 2 O/P
Fixed Attenuators	ad hoc	All	DC - 350 MHz
Impedance Matching Transformers	ad hoc	All	30 - 350 MHz
Balanced Modulator	1	9.2.4	1 kHz - 100 MHz
Waveform Generator	1	9.2.4	rectangular; triangular; $\sin^2$
RF delay line <u>OR</u> Digital Frame Store	1	9.5.3	0 - 20 microseconds

10.4 Human Resources

Until the project has been planned in detail, it is difficult, if not impossible, to come to definitive conclusions as to the staff and person weeks required to carry out the work plan. It is possible however to give, with some confidence, a list of staff and their required capabilities to carry out such a project. These are given below:

- Project Manager - a professional engineer with considerable experience in the field would be required to plan, organize and oversee all summation impairment activities. It is seen that this work would be required on a part-time basis only, and as a first estimate, would possibly entail of an average 25% of a person's time over the period of the project.

- Research Assistant and Subjective Assessment Coordinator - this staff member, seen as a junior professional, would work under the project manager to set up and carry out the subjective assessments. This work would include the procurement, setting up and calibration of all viewing equipment and test equipment; the setting up and maintainence of the viewing room, including making measurements of ambient light and other critical parameters prior to each viewing session; the formulating and drafting of subjective measurement forms; the hiring (on a part-time casual basis) of subjective assessment "viewers"; the briefing and supervision of the "viewers"; analysis of the results under the general direction and supervision of the project manager.

  It is seen that this position would require, as a first estimate, approximately 3/4 of a person's time over the period of the project.

- Viewers - approximately ten viewers would be required for each series of subjective tests. These viewers would be selected from the population in general and hired on a minimum wage basis for the viewing period only. The viewing periods would include a short briefing as well as the actual conduct of the subjective assessment viewing.

11. ADDITIONAL INFORMATION

11.1 FM Noise Versus AM Noise

While this final report was in preparation, work on another project revealed that the subjective impairment of noise in a frequency-modulated channel may be different from that of noise in vestigial sideband amplitude-modulated transmission. If this is true the calculated relationship between C/N and S/N in systems using frequency modulation for all or part of the distribution system may differ from the subjective assessment of noise impairment in otherwise similar situations. In view of the fact that cable television operators use both FM and AM terrestrial microwave systems to carry their signals, measurements to relate the relative impairment due to noise in the two types of modulation are needed. Furthermore, satellite channels are currently always FM channels, hence the need to relate the two for system design and planning. The distinguishing features of FM noise are exhibited above threshold, as well as below. The significance of the difference, i.e. FM versus AM, is not known. It may be slight or significant and it may be sensitive to the absolute magnitude of the impairment.

12. COST EFFECTIVENESS OF IMPAIRMENT UNIT/QUALITY SCALE TECHNIQUES

In attempting to evaluate the cost effectiveness of impairment unit analysis scale techniques for the Department, it is perhaps appropriate to examine the Department's past, current and future work relating to the regulation of, and developmental assistance to, the cable television industry, thereby providing a baseline on which conclusions can be drawn.

The key regulatory document for the cable television industry is Broadcast Procedure 23, "Technical Standards and Procedures for Broadcasting Receiving Undertakings (cable television)". This document defines in some detail the technical standards that cable television systems are required to meet.

To ensure that the standards are met, the Department has three main mechanisms:

- proofs of performance
- informal inspections
- investigations against complaints.

The Proof of Performance is a detailed evaluation of the system, and for the large systems is required to be certified by a staff professional engineer or a consultant. The Proof of Performance is carried out upon start-up and thence every five years or as the Department may decide. The Proof of Performance is carried out under the authority of the appropriate regional office of the Department and all test figures are submitted to this office for approval or disapproval. Headquarters of the Departament retains final authority over approval of applications for, or changes to, cable television undertakings.

Up to approximately four years ago, routine technical inspections were carried out on each cable television system once a year by engineering and technical staff of the Regional Offices of the Department. This work, which included detailed objective testing, involved the use of considerable capital equipment and human resources. It was therefore gradually discontinued to reduce costs. It is planned to replace them with informal visits to discuss licencees' problems and to carry out a subjective assessment of the system. The Department does not at this time have a defined or standardized methodology for subjective assessments, nor does it have guidelines covering the types of impairments to be assessed.

Investigations of individual systems against specific complaints are also carried out. Currently, the normal approach is for a technician from the relevant district office and often an engineer from the regional office to visit the cable company concerned with sophisticated (and expensive) test equipment to check out the system against the subscriber's complaint. On occasion the complaint involves impairments not specified in BP-23 such as direct pick up, co-channel interference or electrical interference. Generally, these impairments are difficult to measure objectively, and unstructured and non-standardized subjective assessments are often made.

An attempt was made to obtain details of capital equipment and human resource expenditures by the Department on cable television regulation enforcement for the five Regional Offices. Unfortunately, the computer program showing these figures was only instituted in 1979, at which time the routine technical inspections had been reduced to an estimated 50%. Based on this estimation some 88 man months of district inspectors' time per year were expended on routine technical inspections. In addition to this must be added an estimated 40 man months of regional engineers' time. Figures for capital equipment investment are not available, but it is estimated that test equipment could be valued in the order of $80,000 to $100,000 for each of the districts, not including cost of vehicles.

In reviewing the validity of the Research Branch of the Department supporting work on this concept, the current long range operational plan of the Department can provide a baseline. In the next section of this Report, conclusions are drawn based upon the following extracts from this operational plan:

- The purpose of communication/information technology (research) is to assist the development of Canadian industries which ..... provide related systems ..... or services and to support the mandate of DOC .....

- To assist industry in developing ..... communications ..... systems aiming to improve the social and economic status of the Canadian people

- To support the Department in its statutory responsibility for ..... regulation of the spectrum

- To ensure liaison with industrial ..... laboratories in Canada ..... on all aspects related to communications, research and development.

To provide a baseline for conclusions on cost-effectiveness within the industry, discussions were held with cable television licencees. These indicated that while subjective assessments of quality are frequently made by maintenance staff, these assessments are uncoordinated and based on a maintenance worker's individual approach to the matter. They are therefore of only limited use in improving cable television service.

For most Canadian cable television companies it is not practical for technicians maintaining the overall distribution system to use large quantities of expensive test equipment for routine maintenance. Diagnostic tests are therefore frequently limited to the use of a signal level meter and a TV set.

The main limitation on improving maintenance practices in the cable industry is one of cost. Any methodology which, whilst keeping cost down could improve upon the evaluation of system performance by maintenance staff with subsequent higher efficiency by repair staff would undoubtedly improve cost-effective within the industry itself. However, experience has shown that despite a strong Association (the CCTA), universal acceptance of a new approach or concept usually only comes after an initiative from the Department.

13. CONCLUSIONS AND COMMENTS

After detailed analysis of the information gained from the literature reviewed during the course of this study, and taking cognizance of the facts and opinions obtained during interviews with researchers and engineers in Canada, United Kingdom and Germany, the following comments are made and conclusions (1) drawn:

- The concept of the derivation of subjective impairment units related in graphical form to objective measurements to permit the repeatable subjective assessments of electronic distribution systems has, over the last 30 years, been the subject of considerable research on a world-wide basis. This research has lead to the acceptance of the concept by the International Radio Consultative Committee (CCIR).

  IT CAN THEREFORE BE CONCLUDED THAT THE GENERAL CONCEPT OF SYSTEM EVALUATION BY PROPERLY RESEARCHED SUBJECTIVE ASSESSMENT IS VALID.

- For broadcasting systems, the use of impairment unit curves, which are directly related to the CCIR five point quality and impairment scales, are used on an operational basis by:

  - The British Broadcasting Company
  - The Independent Television Authority (United Kingdom)
  - Institut Für Rundfunktechnik GMBH. (IRT-West Germany)
  - ORF (Austrian National Radio-Television)
  - The Canadian Broadcasting Corporation

  The majority of these organizations use the impairment unit/quality scale approach for television and radio to evaluate their production, distribution and transmission facilities. In Canada the CBC has found it to be a very cost-effective approach to improving its overall network performance.

  IT CAN THEREFORE BE CONCLUDED THAT THE IMPAIRMENT UNIT/QUALITY SCALE SUBJECTIVE ASSESSMENT CONCEPT HAS BEEN SUFFICIENTLY DEVELOPED TO PERMIT ITS EFFICIENT AND COST-EFFECTIVE USE IN THE BROADCAST INDUSTRY.

---

(1) Conclusions are shown in UPPER CASE.

- The work of the CCIR, CBC and others provides a major base for additional work to permit subjective assessment of Canadian cable television systems.

  IT CAN BE CONCLUDED THAT A COMPARATIVELY MODEST PROGRAM OF WORK, AS DETAILED IN THE BODY OF THIS REPORT, WOULD PERMIT THE USE OF SUBJECTIVE ASSESSMENTS RELATABLE TO OBJECTIVE MEASUREMENTS BY THE DEPARTMENT AND BY INDUSTRY TO EVALUATE CABLE TELEVISION SYSTEMS AGAINST THE REQUIREMENTS OF BP-23.

- A number of impairments typical of cable television systems are difficult to measure objectively and therefore do not figure in BP-23. These include:

  - direct pick-up of a non-phase locked program different to that being carried on the cable television channel
  - co-channel interference
  - electrical interference
  - direct pick-up of a phase locked program different to that being carried on the cable television channel

  They are, however, most amenable to impairment unit evaluation.

  THE USE OF IMPAIRMENT UNIT CURVES TO ASSESS THE PREVIOUS UNMEASURED IMPAIRMENTS TABULATED ABOVE WOULD IMPROVE THE SERVICE PROVIDED BY CABLE TELEVISION SYSTEMS WITHOUT SIGNIFICANT COST INCREASES.

- The study has identified a special case of impairment relating to the alphanumeric programming which frequently makes up a significant portion of the services provided on a cable television system.

  THE DEVELOPMENT AND USE OF IMPAIRMENT UNIT CURVES FOR ALPHANUMERIC PROGRAMMING COULD ALSO IMPROVE SERVICE.

- In attempting to assess the usefulness and cost effectiveness of subjective assessment to the Department of Communications two areas were considered. The first was in relation to the telecommunications regulatory service and their program of routine inspection and investigation of the cable television industry. The second was in relation to the assistance that the Research Branch is committed to provide to the Regulations Branch and to industry under its 1983/1984 operational plan.

- IT CAN BE CONCLUDED FROM THIS WORK THAT THE REGULATIONS BRANCH CAN FULFILL ITS MANDATE TO INSPECT ON A ROUTINE BASIS THE PERFORMANCE OF LICENCED CABLE TELEVISION SYSTEMS BY THE REPLACEMENT OF OBJECTIVE MEASUREMENTS BY A SUBJECTIVE EVALUATION. TO BE EFFECTIVE THIS SUBJECTIVE EVALUATION MUST BE CARRIED OUT BY PERSONNEL TRAINED ON A STANDARD SUBJECTIVE ASSESSMENT VIDEO TAPE PROVIDING A STANDARDIZED METHODOLOGY AND ASSESSMENT VALUES FOR ALL RELEVANT IMPAIRMENTS OF A CABLE TELEVISION SYSTEM. SUCH A PROGRAM WOULD REPLACE EXPENSIVE TEST EQUIPMENT WITH A HIGH GRADE VIDEO PLAYBACK MACHINE AND VIDEO MONITOR. IT IS FURTHER CONCLUDED THAT AN EVALUATION EFFECTIVENESS EQUIVALENT TO THAT OBTAINED BY OBJECTIVE MEASUREMENTS CAN BE REACHED USING PERHAPS ONE THIRD OF THE HUMAN RESOURCES PREVIOUSLY USED.

- IT IS FURTHER CONCLUDED THAT SCIENTIFIC, ENGINEERING AND TECHNICAL ASSISTANCE TO THE REGULATIONS BRANCH AND TO THE INDUSTRY FALLS WITHIN THE MANDATE OF THE RESEARCH BRANCH OF THE DEPARTMENT.

- IT IS CONCLUDED THAT IT WOULD BE MOST COST-EFFECTIVE TO THE DEPARTMENT TO HAVE WORK CARRIED OUT WHICH WOULD MAKE AVAILABLE A COMPLETE SET OF IMPAIRMENT UNIT CURVES FOR NTSC CABLE TELEVISION SYSTEMS. BASED ON THESE CURVES A STANDARD TRAINING TAPE COULD BE PRODUCED FOR USE BY THE REGULATIONS BRANCH OF THE DEPARTMENT.

- IT IS FURTHER CONCLUDED THAT SUCH WORK COULD MOST EFFECTIVELY BE CARRIED OUT BY THE ESTABLISHMENT OF A PERMANENT VIEWING FACILITY IN WHICH THE NECESSARY TESTS FOR THIS AND OTHER SUBJECTIVE ASSESSMENT WORK WOULD BE CARRIED OUT.

- IT IS FURTHER CONCLUDED THAT IF THE DOC SUBJECTIVE ASSESSMENT TRAINING TAPE, WERE USED AS AN INDUSTRY TRAINING INSTRUMENT, TECHNICIANS COULD REPORT DEFECTS IN FAR GREATER DETAIL AND WITH SUFFICIENT ACCURACY TO GIVE SIGNIFICANTLY MORE EFFECTIVE MAINTENANCE AT ESSENTIALLY NO INCREASE IN MAINTENANCE COST.

- THE USE OF THE IMPAIRMENT UNIT/QUALITY SCALE CONCEPT, EVEN IN EVALUATING SINGLE IMPAIRMENTS, WOULD OVERCOME ONE OF THE KEY LIMITATIONS OF OBJECTIVE TESTING, I.E. THE LACK OF A DIRECT RELATIONSHIP TO THE PICTURE VIEWED BY THE SUBSCRIBER.

- AS THE USE OF IMPAIRMENT UNITS FOR THE SUBJECTIVE ASSESSMENT OF CABLE TELEVISION SYSTEMS BECAME GENERAL, IT COULD WELL BE APPROPRIATE FOR THE DEPARTMENT TO SPECIFY THEIR USE FOR AT LEAST PART OF CABLE TELEVISION SYSTEM PROOFS OF PERFORMANCE, THUS INCREASING THE OPERATIONAL AND COST-EFFECTIVENESS OF THESE PROOFS.

- In reviewing work carried out on the summation of impairments and discussing the use of this approach with senior researchers and engineers, IT IS CONCLUDED THAT WHILE THE USE OF THE CONCEPT OF SUMMATION OF IMPAIRMENT UNITS TO PROVIDE AN OVERALL ASSESSMENT OF QUALITY IS CONSIDERED TO BE OF MAJOR IMPORTANCE, WORK IN THIS AREA HAS NOT YET REACHED THE STAGE WHERE IT CAN BE USED OPERATIONALLY WITH FULL CONFIDENCE. IT IS, HOWEVER, USED ON A LIMITED BUT VERY USEFUL BASIS BY THE CBC AND BY OTHER ORGANIZATIONS THROUGHOUT THE WORLD. THERE IS STILL SOME DISAGREEMENT AS TO PRECISE METHODOLOGY PARTICULARLY WHEN THE SUMMATION OF HIGH IMPAIRMENT LEVELS ARE CONSIDERED.

  However, it cannot be over-emphasized that the usefulness of the summation of impairments concept (at an acceptable level of confidence) would be extremely high and would overcome the second limitation of objective testing, the fact that tests do not consider the impact upon the end user of the combination of the impairments present in a distribution system. IT IS CONCLUDED THAT SUMMATION OF IMPAIRMENT UNITS WOULD ALSO BE A POWERFUL TOOL FOR CATV SYSTEM DESIGNERS WITH RESPECT TO BEING ABLE TO TRADE OFF VARIOUS IMPAIRMENTS AND TO ACHIEVE BETTER OPTIMIZED AND HOPEFULLY MORE COST-EFFECTIVE NETWORKS.

- IT CAN ALSO BE CONCLUDED FROM THE WORK CARRIED OUT IN THIS FEASIBILITY STUDY, THAT IT IS HIGHLY LIKELY THAT THE APPROACH CAN BE USED TO ADVANTAGE IN SATELLITE DISTRIBUTION SYSTEMS TO EVALUATE IMPAIRMENTS CAUSED BY CLOSE SATELLITE SPACING, FREQUENCY RE-USE, THE SHARING OF TRANSPONDERS BY MORE THAN ONE CUSTOMER OR SERVICE, ADJACENT CHANNEL CROSSTALK AND THE USE OF NEW MODULATION TECHNIQUES, AMONGST OTHER POSSIBLE IMPAIRMENT PRODUCING MECHANISMS. THERE IS ALSO EVERY INDICATION THAT THE CONCEPT CAN ALSO BE USED TO INCREASE THE COST AND OPERATIONAL EFFECTIVENESS OF EVALUATION OF OTHER ELECTRONIC DISTRIBUTION SYSTEMS SUCH AS TERRESTRIAL MICROWAVE DISTRIBUTION SYSTEMS, AUDIO DISTRIBUTION SYSTEMS AND REMOTE SENSING SYSTEMS. The work carried out to date is in part appropriate to the evaluation of such systems, but it is felt that in each case a detailed review will be required to ascertain additional tests required to use the concept for specific types of distribution systems.

14. RECOMMENDATIONS

From the conclusions drawn in Section 13 above, the following recommendations are made:

- That, after consultation with the cable television industry, the Department of Communications plan work that would result in the availability of curves relating impairment units to objective measurements for all of the cable television impairments defined in Section 5 of this report.

- That to carry out this work a viewing laboratory be established by either the Department, or through agreement with the private sector, to carry out the subjective assessments required for the production of such curves, and that such a viewing laboratory use the approaches defined by CCIR, by other authorities, and within the body of this report.

- That such work be carried out on a step by step basis to minimize initial investment by the Department and by industry.

- That a training tape based on impairment unit curves be produced for Department and cable television industry use.

- That all work be carried out in close cooperation with the cable television industry.

- That further work be considered to advance the state of the art in the use of the summation of impairment units.

- That the concept of the use of impairment units, and the use of the summation of impairment units, be considered in relation to the distribution of audio as well as video signals.

- That work should be undertaken to ascertain the usefulness of the concept of subjective assessment related by curves to objective measurements for other electronic distribution systems, particularly satellite distribution systems.

- If such an approach is considered appropriate, devise a work plan similar to that given in this report, to lead to the production of impairment unit curves for satellite distribution and other electronic distribution systems.

* * *

APPENDIX A

DEFINITIVE DOCUMENTS USED AS A BASIS FOR THIS STUDY

GOVERNMENT OF CANADA
DEPARTMENT OF COMMUNICATIONS

GOUVERNEMENT DU CANADA
MINISTÈRE DES COMMUNICATIONS

BP-23
ISSUE 2

BROADCAST PROCEDURE

(Section 4: Technical Standards Only)

TECHNICAL STANDARDS AND PROCEDURES FOR BROADCASTING RECEIVING UNDERTAKINGS (CABLE TELEVISION)

EFFECTIVE DATE:
JANUARY 1, 1982

TELECOMMUNICATION REGULATORY SERVICE

PR 23
2e ÉDITION

PROCÉDURE SUR LA RADIODIFFUSION

NORMES TECHNIQUES ET MÉTHODES POUR LES ENTREPRISES DE RÉCEPTION DE RADIODIFFUSION (TÉLÉVISION PAR CÂBLE)

MISE EN VIGUEUR:
1er JANVIER, 1982

SERVICE DE LA RÉGLEMENTATION
DES TÉLÉCOMMUNICATIONS

PART 4

TECHNICAL STANDARDS

4.1 General

4.1.1 Unless indicated otherwise, all technical standards herein refer to performance requirements between the cable television system input and any subscriber terminal.

4.1.2 The technical standards herein relate to the measurement procedures of Part 5 of this document and, where appropriate, in the presence of all signals normally carried on the system.

4.2 Television Carrier Levels

4.2.1 The visual carrier level of every television signal at any subscriber terminal shall be not less than 0 dBmV and not more than +14 dBmV.

4.2.2 The visual carrier levels of the television signals on adjacent channels shall not differ from each other by more than 3 dB.

4.2.3 The maximum difference between visual carrier levels within any 90 MHz range of frequencies shall be not greater than 8 dB.

4.2.4 The aural carrier level of every television signal shall be not less than 10 dB and not more than 20 dB below its associated visual carrier level.

4.3 FM Carrier Levels

The carrier level of every FM signal as determined at any subscriber terminal shall be not less than - 14 dBmV and not more than +4 dBmV.

4.4 Carrier to Noise Ratio

The carrier to noise ratio for each television channel at any subscriber terminal shall be not less than 40 dB.

4.5 Carrier to Hum Ratio

The carrier to hum ratio in any television channel at any subscriber terminal shall be not less than 34 dB.

4.6 Crossmodulation Ratio

The crossmodulation ratio in any television channel at any subscriber terminal shall be not less than 48 dB, when the desired modulation is 100% square wave.

4.7 Carrier to Beat Ratio

4.7.1 The carrier to beat ratio for single frequency intermodulation products and other discrete frequency interfering signals in any television channel at any subscriber terminal shall be not less than that indicated by the curve in Figure 4.1.

4.7.2 Clause 4.7.1 is not applicable to intermodulation products or other discrete signals which are zero-beat with visual carriers.

4.7.3 Within the FM band, the products and signals referred to in clause 4.7.1 shall be at least 30 dB below the minimum FM carrier level.

4.7.4 In other bands, the products and signals referred to in clause 4.7.1 shall be at least 30 dB below the level of the nearest television visual carrier.

4.8 Carrier to Composite Beat Ratio

4.8.1 The carrier to composite beat ratio in any television channel at any subscriber terminal shall be not less than 53 dB.

4.8.2 Clause 4.8.1 is not applicable to intermodulation products which are zero-beat with visual carriers.

4.9 Echo Rating

The echo rating in any television channel at any subscriber terminal shall not exceed 7%.

4.10 Frequency Response (Broadband System)

The frequency response of the broadband system for each television channel that is distributed shall be within ± 1 dB over the 6 MHz band of frequencies of the television channel.

4.11 Chrominance-Luminance Delay Inequalities (Broadband System)

The chroma delay introduced by the broadband system and measured at any subscriber terminal shall be within ± 150 nanoseconds.

4.12 Radiation

Radiation from a cable television system shall not exceed the limits set out in the following table:

Frequency Range (MHz)	Radiation Limit (microvolts per metre)	Test Distance (metres)
up to 54	20	10
54 - 108	20	3
108 - 174	10	3
174 - 216	20	3
216 - 444	20	10

4.13 Frequency Stability of Headend Equipment

4.13.1 The frequency stability of headend equipment shall be such that visual and aural carrier frequencies of television signals are maintained within 25 kHz of assigned carrier frequencies, under all environmental conditions and power supply variations.

4.13.2 The frequency separation between a visual carrier and the associated aural carrier of any television signal distributed by a cable television system shall be maintained at 4.5 MHz $\pm$ 2 kHz.

4.13.3 The frequency stability of headend equipment shall be such that FM carrier frequencies are maintained within 10 kHz of assigned carrier frequencies, under all environmental conditions and power supply variations.

4.13.4 The frequency stability of headend equipment shall be such that pilot carrier frequencies are maintained within 25 kHz of assigned carrier frequencies, under all environmental conditions and power supply variations.

4.14 Differential Gain and Phase of Modulators (Headend)

4.14.1 The differential gain for the 3.58 MHz colour subcarrier of a television signal shall not exceed 2 dB or 26%.

4.14.2 The differential phase of the 3.58 MHz colour subcarrier of a television signal shall not exceed $\pm$ 5 degrees.

4.15 Frequency Response of Processors (Headend)

The RF frequency response for each television channel signal processor shall, when the channel is swept with a constant amplitude signal across the frequency range from vcf-0.5 MHz to vcf+4 MHz with the response at vcf + 0.2 MHz used as a reference value, be:

(1) within ± 1.5 dB of the reference value, across the frequency range from vcf - 0.5 MHz to vcf + 3.58 MHz; and,

(2) between +1 and -2 dB of the reference value, at vcf + 4 MHz.

4.16 Frequency Response of Modulators (Headend)

The RF frequency response for each television video modulator shall, when the information portion of an input video signal is a constant amplitude sine wave with the response at vcf + 0.2 MHz used as a reference value, be:

(1) within ± 1.5 dB of the reference value across the frequency range vcf - 0.5 MHz to vcf + 3.58 MHz; and,

(2) between + 1 and - 4 dB of the reference value, at vcf - 0.75 MHz and at vcf + 4 MHz; and,

(3) 20 dB or greater below the reference value, at vcf - 1.5 MHz.

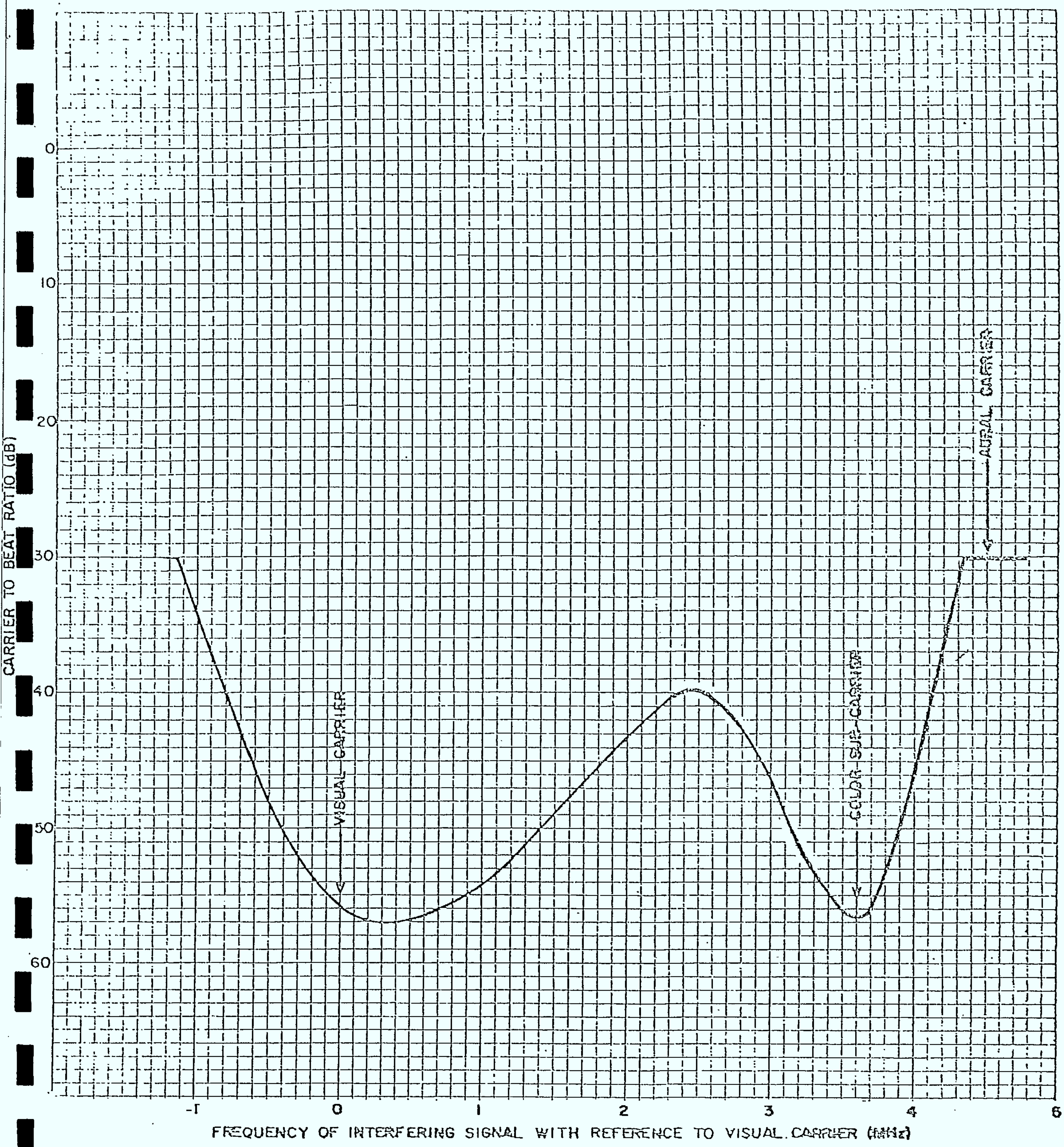

FIGURE 4·1 PERMISSIBLE LIMITS FOR MINIMUM CARRIER TO BEAT RATIO (SINGLE FREQUENCY)

CHAPTER III: PERFORMANCE REQUIREMENTS

SECTION SEVEN – PERFORMANCE REQUIREMENTS FOR SYSTEMS OPERATING BETWEEN 30 MHz AND 1 GHz

19 Object

The object of this chapter is to define system performance limits which will, with an unimpaired input, produce picture and sound signals where the impairment to any single parameter will be not worse than Grade four on the five-grade impairment scale contained in CCIR Recommendation 500-1 (Kyoto, 1978, Vol. XI) as given below:

5 – Imperceptible.
4 – Perceptible but not annoying.
3 – Slightly annoying.
2 – Annoying.
1 – Very annoying.

20. General requirements

20.1 All requirements refer to the performance limits which shall be obtained between the input(s) to the head end or head ends and any system outlet when terminated in a resistance equal to the nominal load impedance of the system, unless otherwise specified. Where system outlets are not used, the above applies at the subscriber's end of the subscriber's feeder. In this chapter all references to "system outlet" shall also apply to this case.

20.2 The performance limits set out in this chapter apply when the methods of measurement given in Chapter Two are employed, and, where appropriate, in the presence of all the signals for which the system was designed.

20.3 The performance limits shall be met for those specified conditions of temperature, humidity, mains supply voltage and frequency which apply for the location in which the system is situated, such conditions being determined by each local standards authority.

20.4 The requirements for satisfactory data signal transmission are dealt with in Clause 40.

21. Impedance

The nominal impedance of the system shall be 75 Ω. It should be noted this value applies to all coaxial feeder cable and system outlets and should be used as the reference impedance in level measurements.

22. Carrier levels at system outlets

22.1 *Minimum and maximum carrier levels*

The minimum and maximum carrier levels will depend on many factors, including the performance of typical receivers in use and local installation practices. These local needs are therefore not specified here. Each country should determine which levels are appropriate to the prevailing circumstances. Notwithstanding the above, the maximum levels shall not exceed, and the minimum levels shall be not less than those shown in Table II. It should be noted that on systems carrying large numbers of television channels (say more than 12), it may be necessary to reduce the maximum level to below that given in Table II.

TABLE II

Carrier signal levels at system outlets

Frequency range and service	Maximum level (dBμV)	Minimum level (dBμV)	Definitions
30 MHz to 300 MHz television	83	57	These levels are expressed as the r.m.s. voltage of each vision carrier at the peak of the modulation envelope when measured at the system television outlet across an external 75 Ω termination or referred to 75 Ω.
300 MHz to 1 000 MHz television	83	60	
F.M. sound V.H.F. band mono	80	37	These levels are expressed as the r.m.s. value of each f.m. carrier when measured at the system f.m. outlet across an external 75 Ω termination or referred to 75 Ω.
F.M. sound V.H.F. band stereo	80	47	

Notes 1. – The system should be designed to provide the levels appropriate to stereophonic signals unless it is known that these will never be distributed. Where successive f.m. channels are at an interval of 300 kHz it would be desirable not to exceed 66 dB/μV, and where the interval is 400 kHz, not to exceed 74 dB/μV.

2. – The level of the television sound carrier relative to that of the vision carrier which is permissible varies according to the television system and performance of typical receivers, and therefore each country should determine appropriate limits for relative levels.

22.2 *Carrier level differences*

The differences in carrier levels shall not exceed the values given in Table III.

If f.m. sound signals are present at the system outlet intended for television signals, the level of any f.m. carrier shall be at least 3 dB lower than the lowest television signal level at the outlet.

TABLE III

Maximum level differences at system outlets between distributed television channels

Frequency ranges	Maximum level difference (dB)
30 MHz to 1 000 MHz	15
30 MHz to 300 MHz	12
Any 60 MHz range in v.h.f.	8
Any 100 MHz range in u.h.f.	9
Adjacent channel	3

23. Mutual isolation between system outlets

23.1 *General*

The isolation at any signal frequency between two system outlets (see Sub-clause 20.1) connected separately to a spur feeder via separate subscribers' feeders or looped system outlets shall be not less than 22 dB. This value assumes that the frequency allocations of the distributed channels have been planned to avoid interference from local oscillators of television and f.m. receivers having regard to the intermediate frequencies used in those receivers.

Notes 1. – In some countries a higher value of minimum isolation has been found necessary.

2. – The isolation between any television outlet and any radio outlet should be not less than 60 dB in the frequency ranges 150 kHz to 1 605 kHz, when signals are distributed in that range.

23.2 *Additional requirements if unwanted frequencies are unavoidable*

When the channel allocations or channel conversions are such that the television or f.m. receivers' local oscillator fundamental or harmonic frequencies fall in the f.m. or television channels, the isolation at any signal frequency between two system outlets connected to a spur feed via separate subscribers' feeders or looped system outlets shall be at least 46 dB.

Where the local oscillator signals fall in an f.m. channel, the signal level of the f.m. channel at the system outlet shall be at least 54 dB/μV.

The requirements stated here may not be appropriate in specific national cases, for example where television or f.m. signals are distributed on cabled distribution systems on frequencies outside those designated for broadcast transmission. Where these cases occur, the national authorities are advised to define appropriate isolation limits.

24. Frequency response within a television channel at any system outlet (see Sub-clause 20.1)

24.1 *Amplitude response*

The amplitude response as a function of frequency for the entire system shall be such that the variation in gain over any television channel (bandwidth appropriate to the television system in use) is not more than ± 2 dB relative to that at the vision carrier frequency, and the gain shall not vary by more than 0.5 dB within any frequency range of 0.5 MHz.

Note. – Reception difficulties may impose selectivity requirements on head end equipment which may cause these limits to be exceeded.

24.2 *Phase response*

When the limits defined in Clauses 35 and 39 are met, the requirements for phase response are also met. No specific or additional recommendations for phase response are therefore required.

25. Frequency stability of distributed carrier signals

Where a signal is not distributed at the received frequency or is locally generated, the variation of frequency from the declared nominal due to the system equipment shall not exceed ±75 kHz for a television signal or ±12 kHz for an f.m. sound signal. Where the system carrier frequencies are generated locally the frequency difference between vision and sound carriers for any one channel shall be maintained within ±5 kHz of the nominal for the television system in use (see special remark for System M hereinunder). When adjacent television channels are used, the frequency variation of each of the vision carriers shall not exceed ±20 kHz.

For signals in areas using the System M transmission system, the frequency difference between vision and sound carriers shall be maintained at 4.5 MHz ± 2 kHz.

26. Generation of spurious signals

Frequency converters shall conform to C.I.S.P.R. Publication 13: Limits and Methods of Measurement of Radio Interference Characteristics of Sound and Television Receivers, for television receivers in respect of the level of the r.f. voltage produced at their signal terminals at the fundamental and harmonic frequencies of their local oscillators.

Note. – Where the local oscillator frequencies and harmonics are such that interference to the distributed frequencies is possible, additional measures to reduce unwanted r.f. voltages may be necessary.

27. Timebase and video frequency voltages injected into the mains supply

Equipment shall comply with the limits specified in C.I.S.P.R. Publication 13 for television receivers in respect of r.f. voltages injected into the mains.

28. Intermediate frequency interference

At any system outlet, the level of any signal in the i.f. range of the television receivers shall be more than 10 dB lower than the lowest v.h.f. television signal level and not higher than the lowest u.h.f. television signal level.

29. Random noise

The carrier to noise ratio for systems from the head end input to the system outlets (see Sub-clause 20.1) shall be not less than the values shown in Table IV.

This carrier to noise ratio should be obtained with a test signal applied at the system input equal in level to that normally available at that point, except where the normal input is less than the minimum shown in Table II, in which case the minimum levels given in that table should be used.

TABLE IV

System	Minimum carrier to noise ratio (dB)	Noise bandwidth (MHz)
405-lines system A	40	2.75
625-lines system I	43	5.08
B, C, G and H	43	4.75
L and K_1	43	5.58
D, K	43	5.75
525-lines system M	42	3.33
819-lines system E	46	9.23
FM sound (mono)	41	0.20
FM sound (stereo) (see Note 2)	51	0.20

Notes 1. – Carrier to noise ratio expressed in decibels is defined as:

$$\frac{C}{N} = 20 \log \frac{\text{carrier voltage}}{\text{noise voltage}}$$

Where the carrier voltage is the r.m.s. value of the vision carrier at the peak of the modulation envelope or the r.m.s. value of the f.m. sound carrier, and the noise voltage is the r.m.s. value of the random noise in the channel.

2. – This value applies when the level at the system outlet is the minimum given in Table II.

30. Single-frequency interference to television channels

This clause refers to single-frequency interference which may result from intermodulation or the presence of interfering signals.

At any system outlet the level of any unwanted signal generated within the system shall be such that the lowest carrier to interference ratio within a wanted television channel shall be not less than 57 dB where this ratio is expressed as:

$$20 \log \frac{\text{r.m.s. vision carrier signal voltage}}{\text{r.m.s. of interference voltage}}$$

and the voltages have those values which occur at the peaks of the modulation envelopes. However, where a frequency assignment taking account of known future off-air and distributed channels is adopted so that interference signals fall only in the less sensitive areas of the television channel spectra, a limit lower than that given above may be acceptable (see curves given in Figures 23a to 23e, pages 95 to 99).

31. Single-channel intermodulation interference in television channels

In this special case of single-frequency interference when measured by the method defined in Sub-clauses 8.4 and A2.2, the ratio of the reference level relative to the interference signal shall be not less than 54 dB.

32. Multiple-frequency intermodulation interference to television channels

Under consideration.

33. Cross-modulation between television channels

Under consideration.

34. Differential gain and phase in television channels

The differential gain and phase in any television channel shall not exceed the figures given in Table V.

TABLE V

System	Maximum differential gain (%)	Maximum differential phase (degrees)
NTSC	10	5
PAL	10	12
SECAM	40	32

35. Echoes in television channels

The echo rating as determined at any system outlet when measured by the method defined in Clause 14 shall not exceed 7%.

36. Hum modulation of carriers in television channels

For those systems using other than system M transmission:

At any system outlet the spurious modulation of any vision carrier at the frequency of the supply mains and harmonics thereof shall be such that the reference modulation to hum modulation ratio is not less than 46 dB, and of any a.m. sound carrier shall be such that this ratio is not less than 60 dB where this is expressed as:

$$20 \log \frac{\text{reference modulation}}{\text{peak-to-peak hum modulation}}$$

When the reference modulation is a vision signal, its amplitude is that of the peak-to-peak composite signal, from peak white to sync tip. For an a.m. sound signal, the reference modulation depth is taken as 40%.

For those using system M transmission, the vision carrier ratio of reference modulation to hum modulation shall be not less than 35 dB.

37. Radiation

37.1 *Radiation from complete systems*

Due to the extreme variability of field conditions, repeatable measurements of radiation from operating systems are impracticable, and absolute performance requirements are inappropriate.

37.2 *Radiation from individual system components*

Under consideration.

38. Immunity to external fields

38.1 *Immunity of complete systems*

The immunity of the system shall be such that at any system outlet (see Sub-clause 20.1) on any distributed channel, the ratio of carrier to interfering signal (caused by an external field) shall be not less than the limits given for single-frequency interference in Clause 30 and Sub-clause 41.5.

38.2 *Immunity of individual system components*

Under consideration.

39. Chrominance/luminance delay inequality

At any system outlet (see Sub-clause 20.1) on any television channel the difference in transmission delay between luminance and chrominance information shall not exceed 100 ns.

40. Requirements for data signal transmission

Under consideration.

41. F.M. radio: Additional performance requirements

41.1 *Amplitude response within an f.m. channel*

The amplitude response as a function of frequency for the entire system shall be such that the maximum amplitude variation over any f.m. channel (bandwidth appropriate for the transmission system in use) is not more than 3 dB with the slope not exceeding 0.3 dB per 10 kHz within 75 kHz of the carrier.

41.2 *Phase response within an f.m. channel*

Under consideration.

41.3 *Adjacent channel spacing*

The minimum spacing between adjacent unmodulated carriers shall be not less than 400 kHz for high fidelity transmission and not less than 300 kHz for other f.m. services.

41.4 *Relative level of adjacent carriers*

The level difference between any two carriers in the v.h.f. band allocated to f.m. broadcasting shall not exceed 8 dB. The level difference between carriers on adjacent channels with less than 600 kHz spacing shall not exceed 6 dB.

41.5 *Single-frequency interference within an f.m. channel*

Under consideration.

41.6 *Multiple-frequency interference within an f.m. channel*

Under consideration.

SECTION EIGHT – SAFETY

42. Safety requirements

A cabled distribution system shall be so designed, constructed and installed as to present no danger, either in normal use or under fault conditions, to subscribers, to personnel working on, or externally inspecting the system, or to any other person, providing particularly:

CCIR
XVth Plenary Assembly
Geneva, 1982

Doc. 11/1013-E
2 November 1981

Ref. : Doc. 11/5038 + Corr.1

DRAFT

RECOMMENDATION 500-1 (MOD F)

METHOD FOR THE SUBJECTIVE ASSESSMENT OF THE QUALITY OF TELEVISION PICTURES

(1974 – 1978)

The CCIR,

CONSIDERING

(a) that a large amount of information has been collected about the methods used in various laboratories for the assessment of picture quality;

(b) that examination of these methods shows that there exists a considerable measure of agreement between the different laboratories about a number of aspects of the tests;

(c) that the adoption of a single standardized method is of the greatest importance in the exchange of information between various laboratories;

(d) that routine or operational assessments of picture quality and/or impairments using a five-grade quality and impairment scale made during the normal course of duty of certain supervisory engineers, can also make some use of certain aspects of the methods recommended for laboratory assessments,

RECOMMENDS

that the general methods of test, the grading scales and the viewing conditions for the assessment of picture quality, described in Annex I should be used for laboratory experiments and for operational assessments whenever possible.

ANNEX I

METHOD FOR LABORATORY ASSESSMENT OF PICTURE QUALITY

1. Introduction

Before giving details of the method, it should be remarked that the ultimate purpose of tests is to discover the acceptability of some impairment whose complete elimination may be uneconomic to achieve. Although in a normal television audience there will be some "expert observers" *, the proportion of them is thought to be so small that it is proper to concentrate the objective of the laboratory tests on the opinions of non-experts, because the use of experts could lead to results which are much more critical than would be obtained with normal television viewers. The philosophy which leads to preference being given to the opinions of non-experts is also applied to the choice of test pictures and viewing conditions which are chosen to be more critical than average but not unduly so. As tests with non-expert observers tend to be lengthy, it is often desirable that a quick test should be carried out by experts. In this case, a smaller number of observers can be used. However, it should be noted that in certain circumstances tests carried out with expert observers may not be a satisfactory substitute for tests carried out by non-expert observers [CCIR, 1963-66]. In cases of doubt, the relationship between expert and non-expert opinion should be investigated.

The statistical design of experiments has been well considered and documented; the amount of data which needs to be collected depends upon such interrelated factors as the confidence level which is needed in the answer, the standard deviation in the measurements, and the relative magnitude of the effect which it is required to detect. The following suggestions are intended as guide lines to assist in formulating a considered experimental assessment programme.**

2. Details of the method

2.1 *Observers*

Observers may be either expert or non-expert.

Normally, the experiments should be carried out with non-expert observers (see Introduction). These observers should be appropriately introduced to the test method, the grading scale and the impairments or processes to be assessed. Remarks likely to influence the rating should be avoided.

The number of observers should be at least ten, and preferably about twenty. In all cases, the number and type of the observers, as well as the details of their introduction to the test should be specified.

2.2 *Grading scales*

The five-grade scales are listed in Table I. It is important to restrict the definitions of the grades to those given in Table I; according to the nature of the problem it may be more appropriate to use a quality scale or an impairment scale.

TABLE I

Five-grade scale	
Quality	Impairment
5 Excellent 4 Good 3 Fair 2 Poor 1 Bad	5 Imperceptible 4 Perceptible, but not annoying 3 Slightly annoying 2 Annoying 1 Very annoying

* The term "expert observers" is considered to apply to observers who have had recent extensive experience in observing picture quality or impairments, particularly of the type being studied in the subjective test.

** Draft Report 405-3 (MOD F) gives detailed information on the application of these guiding principles. Draft Recommendation (AA/10-11) also contains some additional information relating to specific applications.

It is recommended that results should be presented in the form given above although it is recognised that, to suit local practices, some laboratories may wish to invert the order of the numbering or to replace the numbers with letters when an experiment is being carried out.

The scale which has been used, that is, an impairment or a quality scale, should always be quoted along with the results of an experiment. *

For certain types of experiment, a comparison scale is more convenient than a quality or impairment scale; in such cases the scale of Table II is recommended:

TABLE II *Comparison scale*

+3	Much better
+2	Better
+1	Slightly better
0	The same
−1	Slightly worse
−2	Worse
−3	Much worse

2.3 Test pictures

About five still pictures and/or sequences including moving subject matter, as appropriate, should be used. Normally, these should be pictures that are more critical than average pictures, but not unduly so, taking into account the specific assessements being made. For example, certain effects are best evaluated by using scenes containing bright saturated colours viewed by colour cameras. Other effects may dictate the use of sharp, contrasted edges, detailed areas and/or de-saturated colours with a low level of brightness. In assessments where the effects of movement do not cause specific impairments the use of still pictures should be preferred. However, test patterns should not normally be included. The visual scenes and the signal sources used should be described in the presentation of the results.

2.4 *Viewing conditions*

The preferred viewing conditions are affected by the field frequency of the television system. Table III shows the conditions for systems with 50 and 60 fields/s.

---

* In view of the large number of documented results which have been obtained using six-grade scales, it is desirable to have a means of transforming results so that these data can still be used. Uncertainties arise in attempting to transform results obtained from one scale into another scale, particularly from the categorisation of the grades. However, as a first approximation, the following linear relationship can be used to transform a mean value A_6 obtained in an experiment using a six-grade scale (Report 405-3, Notes 7 and 9) into a corresponding mean value A_5 in the five-grade scale:

$$A_5 = 5.8 - 0.8\,A_6.$$

TABLE III

Condition	Field frequency	50 fields/s	60 fields/s
a	Ratio of viewing distance to picture height	6 (1)	4 to 6
b	Peak luminance on the screen (cd/m^2)	70 ± 10 (2)	70 ± 10
c	Ratio of luminance of inactive tube screen (beams cut off) to peak luminance	<0·02	<0·02
d	Ratio of the luminance of the screen when displaying only black level in a completely dark room, to that corresponding to peak white	approximately 0·01	
e	Ratio of luminance of background behind picture monitor to peak luminance of picture	approximately 0·1 (3)	approximately 0·15
f	Other room illumination	low (4)	low
g	Chromaticity of background	white (5)	D_{65}
h	Ratio of solid angle subtended by that part of the background which satisfies this specification, to that subtended by the picture	>9	

(1) Normally 6; if a different ratio is used, this should be stated.

(2) Normally (70 ± 10) cd/m^2, or (220 ± 30) asb (6), but certain tests may require luminances outside the tolerances, for example, because of flicker, defocusing, etc.

(3) If the ratio is greater than 0·1, the chromaticity has to be nearer to Illuminant D_{65} (7).

(4) The specification is loosely phrased here, since the precise value is not critical, provided it does not conflict with condition *c*.

(5) Not very critical. Any white in the region between standardized illuminants A and D_{65}. See, however, Note (3).

(6) 1 apostilb (asb) $= \frac{1}{\pi}$ candela per square metre (cd/m^2).

(7) Illuminants standardized by the International Commission on Illumination (CIE); see International Electrotechnical Vocabulary, Group 45, No. 45-15-145.

2.5 *Presentation*

The pictures and impairments should be presented in random sequence with the proviso that the same picture (i.e. test scene or sequence) should never be presented on two successive occasions with the same or different levels of impairment.

When using the quality or impairment scale, the range of impairments should be chosen, wherever practicable, so that all grades are used by the majority of observers; a grand mean score (averaged over all judgements made in the experiment) close to 3 should be aimed at, to standardize results.

A session should not last more than roughly half an hour, including the explanations and preliminaries; the test sequence could begin with a few pictures indicative of the range of impairments; judgements of these pictures would not be taken into account in the final results.

REFERENCES

CCIR Document

[1963-66]: XI/16, United Kingdom (refers to a published paper: PROSSER, R. D., ALLNATT, J. W. and LEWIS, N. W. [March, 1964] Quality grading of impaired television pictures. *Proc. IEE*, 111, 491-502).

BIBLIOGRAPHY

BERNATH, K., KRETZ, F. and WOOD, D., [April 1981], The EBU method for organizing subjective tests of television picture quality. EBU Review-Technical, No. 186.

CCIR Documents

[1966-69]: XI/158 (France).

[1970-74]: 11/111 (Italy); 11/128 (U.S.S.R.).

CCIR
XVth Plenary Assembly
Geneva, 1982

Doc. 11/1015-E
11 November 1981

Ref. : Doc. 11/5047

DRAFT

REPORT 405-3 (MOD F)

SUBJECTIVE ASSESSMENT OF THE QUALITY OF TELEVISION PICTURES

(Study Programmes 3A-1/11 and 3B/11)

(1966-1970-1974-1978)

Recommendation 500-1 proposes a method for subjective assessment of the quality of television pictures and provides a partial answer to Study Programme 3A-1/11. However, as many results of subjective experiments which have been carried out in different countries using different methods are already available and in use, it is considered important that the details of the grading scales and methods of test employed should remain on record. This information is summarized in Annex I.

In order to facilitate the comparison of results among different laboratories it would be useful to list the types of experiments for which the different scales (that is quality, impairment and comparison) are best used.

Study Programme 3A-1/11 calls for studies on the analysis and presentation of the results of subjective tests. Contributions have been received on this topic from a number of Administrations and Annex II lists and summarizes these contributions.

Annex III is an example of one such method. Administrations are invited to provide further contributions on all aspects of this important topic.

Annex IV describes a number of procedures using either the quality scale, or the impairment scale, and includes a comparison of the results obtained with a number of different methods and a discussion of their relative merits.

Annex V is a first attempt to describe methods for testing the performance of digital television systems.

Annex VI describes an initial tentative description of the methods of measuring the subjective quality of alphanumeric and graphic pictures.

Annex VII concerns the specifications and alignment procedures for signal sources and displays.

BIBLIOGRAPHY

ALLNATT, J. W. [1965] Subjective quality of colour-television pictures impaired by gain and delay inequalities between the luminance and chrominance channels. *Proc. IEE*, 122, 1819-1824.

ALLNATT, J. W. [1971] Subjective assessment of impairments in television pictures. *Applied Ergonomics*, 2, 231-235.

ALLNATT, J. W. [1975] Opinion-rating model applied to television-transmission-system studies. *Proc. IEE*, 122, 769-774.

ALLNATT, J. W. and CORBETT, J. M. [1972] Adaptation in observers during television quality-grading tests. I, Adaptation as a function of the conditioning situation. *Ergonomics*, 15, 353-366.

ALLNATT, J. W. and CORBETT, J. M. [1972] Adaptation in observers during television quality-grading tests. II, Progress of adaptation during an experiment, *Ergonomics*, 15, 491-504.

ALLNATT, J. W. and PROSSER, R. D. [1966] Subjective quality of colour-television pictures impaired by random noise, *Proc. IEE*, 113, 551-557.

CORBETT, J. M. [1970] Effect of observer adaptation on the results of television quality-grading tests, *Proc. IEE*, 117, 512-514.

CORBETT, J. M. and ALLNATT, J. W. [1974] Subjective quality of television pictures impaired by short-delay echo, *Proc. IEE*, 121, 583-593.

SIOCOS, C. A. [1971] Numerical values for subjective picture quality in television. *Journal of the SMPTE*, 80, 958-962.

WEAVER, L. E. [1968] The quality rating of colour television pictures. *Journal of the SMPTE*, 77, 601-612.

CCIR Documents

[1966-69]: XI/158 (France); XI/159 (France);

[1970-74]: 11/89 (France); 11/26 (Canada); 11/111 (Italy).

TABLE I – *Conditions for laboratory assessments*

Reference	UK [CCIR 1963-66a]	EBU/OIRT [CCIR 1963-66c and d]	Fed. Rep. of Germany [CCIR 1963-66b]	Japan [CCIR 1963-66e and 1966-69a]	USA [CCIR 1963-66f]	USA [CCIR 1966-69b]	Italy [CCIR 1966-69c]	People's Republic of China [CCIR 1978-82]
Observers Category Number	Non-expert 20-25	>6	Non-expert >10	Non-expert 20-25	Non-expert approx. 200	Expert >10	Non-expert approx. 20	Non-expert 250 expert 65
Grading scale Type Number of grades	Quality 5 (Note 1)	Impairment 6 (Note 9) Quality 6 (Note 7) Comparison 7 (Note 11)	Impairment 5 Quality 5 Comparison 5	Impairment 5 (Note 3) Quality 5 (Note 2)	Quality 6 (Note 4)	Impairment 7 (Note 10)	Comparison 5 (Note 5)	Quality 5
Test pictures Number	4-8	5	>5	>3	2-8	3-4	6	5
Viewing conditions Ratio of viewing distance to picture height	6	4-6	6	6-8	6-8	4	6	4-10
Peak luminance on the screen (cd/m²) [2]	50	41-54	50	Approx. 400 (monochrome) 74-84 (colour) (Note 12)	70 (Note 12)	170 (monochrome) 34 (colour) (Note 12)	50	85
Contrast range of the picture	Not specified		Not specified	30/1 to 50/1	Not specified			
Luminance of inactive tube screen (cd/m²)	<0·5	0·5	<0·5	Approx. 5 (monochrome) 0·7-2 (colour)	2		Approx. 0·5	
Luminance of backcloth (cd/m²) [3]	1 Illuminant C		Approx. 2·5 [1]	Not specified				
Room illumination, average (lx)	3			Not specified	6·5			
Presentation	Random sequence of pictures and impairments		Random sequence of pictures and impairments	Random sequence of pictures and impairments	Random sequence of impairments	Random sequence of impairments	Random sequence of pictures and impairments	Random sequence of pictures and impairments

(¹) For monochrome only. (²) 1 cd/m² = 1 nt = 3·14 asb = 0·29 ft-L. (³) Ambient luminance at the end of the room as seen by the viewer.

TABLE II – *Conditions for assessment during programme transmission*

Reference	OIRT [CCIR, 1966-69d]		Canada [CCIR, 1966-69e]
Observers			
Category Number	Expert 1 or 2		Expert 1 or 2
Grading scale			
Type Number of grades	Impairment 6 (Note 9)	Quality 6 (Note 7)	Impairment 5 (Note 4)
Pictures			
Type	Television programmes		Television programmes
Viewing conditions			
Ratio of viewing distance to picture height	4-6		4-6
Angle of view, from a line normal to the face of the monitor			$<30°$
Luminance, on the screen, at reference white (cd/m²)			70 ± 7
Chromaticity of the screen at reference white			Illuminant *D*
Luminance of the inactive tube screen	Adapted to the ambient illumination		As low as practicable
Luminance of "light surround" (cd/m²)			$10{\cdot}5 \pm 3{\cdot}5$ (Note 14)
Chromaticity of "light surround"			Illuminant *D*

Five-grade scales

Quality scales

Note 1. – A Excellent
B Good
C Fair
D Poor
E Bad

Note 2. – 5 Excellent
4 Good
3 Fair
2 Bad
1 Very bad

Impairment scales

Note 3. – 5 Imperceptible
4 Perceptible but not annoying
3 Somewhat annoying
2 Severely annoying
1 Unusable

Note 4. – 1 Imperceptible (implied grade)
2 Detectable
3 Noticeable
4 Objectionable
5 Unsuitable for broadcast

Comparison (fidelity) scales

Note 5. – 1 Equal
2 Slightly different
3 Different
4 Very different
5 Extremely different

Note 6. – +2 Much better
+1 Better
0 The same
–1 Worse
–2 Much worse

Six-grade scales

Quality scales

Note 7. — 1 Excellent
2 Good
3 Fairly good
4 Rather poor
5 Poor
6 Very poor

Note 8. — 1 Excellent: the picture is of extremely high quality, as good as you could desire.
2 Fine: the picture is of high quality providing enjoyable viewing; interference is perceptible.
3 Passable: the picture is of acceptable quality; interference is not objectionable.
4 Marginal: the picture is poor in quality and you wish you could improve it; interference is somewhat objectionable.
5 Inferior: the picture is very poor but you could watch it; definitely objectionable interference is present.
6 Unusable: the picture is so bad that you could not watch it.

Impairment scale

Note 9. — 1 Imperceptible
2 Just perceptible
3 Definitely perceptible but not disturbing
4 Somewhat objectionable
5 Definitely objectionable
6 Unusable

Seven-grade scales

Impairment scale

Note 10. — 1 Not perceptible
2 Just perceptible
3 Definitely perceptible, but only slight impairment to picture
4 Impairment to picture, but not objectionable
5 Somewhat objectionable
6 Definitely objectionable
7 Extremely objectionable

Comparison scale

Note 11. — −3 Much worse
−2 Worse
−1 Slightly worse
0 Same as
+1 Slightly better
+2 Better
+3 Much better

Note 12. — These higher values of peak luminance can be used with 60-field systems.

Note 13. — The Doc., [CCIR, 1963-66g], points out the need for subjective assessment of picture quality during an international programme exchange. The document provides instructions on how to carry out the assessment, and also a list of terms related to the parameters subjected to assessment. The operational experience gained during transmissions in the Intervision network shows that in spite of the variety of monitoring devices used, adequate agreement on the assessment of picture quality was achieved by the technical staff.

Note 14. – The expression "light surround" is defined as the light visible to the observer from a plane or from behind a plane coincident with, and surrounding but not including, the viewing screen. It is provided over an area at least eight times the area of the monitor screen, or, in the case of adjacent monitors, over an area at least four times the total monitor screen area.

REFERENCES

CCIR Documents

[1963-66]: a. XI/16 (United Kingdom); b. XI/31 (Federal Republic of Germany); c. XI/33 (EBU); d. XI/45 (CCIR Secretariat); e. XI/140 (Japan); f. XI/145 (United States of America); g. XI/149 (OIRT).

[1966-69]: a. XI/155 (Japan); b. XI/8 (United States of America); c. XI/45 + XI/181 (Italy); d. XI/46 (OIRT); e. XI/146 (Canada).

/1978-82/: 11/132 (People's Republic of China).

ANNEX II

ANALYSIS AND PRESENTATION OF THE RESULTS OF TELEVISION SUBJECTIVE TESTS

1. Introduction

Methods of analyzing and presenting experimental data relating the degradation of a television signal to the consequent picture impairment, assessed subjectively, have up to now varied considerably among experimenters. Such differences add to the difficulty of comparison of results which may already exist due to variations in experimental method. Generally, the techniques that have been used fall into the four following categories:

1.1 In what is probably the simplest method, numbers in an arithmetic series are assigned to the grades of the assessment scale, thus making possible the computation of a mean subjective score (from the reactions of all observers) corresponding to a given magnitude of objective impairment. The results are usually presented as a smooth-curve plot of mean score against objective magnitude. Statistically, the mean is the most efficient estimator of central tendency of the data, but it gives no information about the distribution of opinions regarding a given impairment condition.

1.2 In a second method [Cavanaugh and Lessman, 1971], analysis is made of the cumulative proportions of observers judging the subjective quality of a given picture condition as being in or above a stated grade. The results are presented as a family of cumulative probability curves for the various grades. The large amount of detail presented tends to make interpretation difficult.

1.3 A third method [CCIR, 1966-69a] employs mean score, as described for the method of § 1.1, but provides summarized information about opinion distributions by quoting the standard deviation of scores about the mean. The method works well for mean scores in the vicinity of mid-scale, but is complicated by the inevitable reduction in standard deviation for low and high values of mean, due to the bounded nature of the scales and the associated phenomenon known as "skewing".

1.4 This difficulty in the method of § 1.3 is overcome in a fourth method [Prosser *et al.*, 1964]. Here again, the method of § 1.1 is relied upon to describe the central tendency of opinions. Use is made of the assumption that the distributions of opinions for a given type of impairment are matched by a theoretical model which can be readily derived from the binomial distribution. Thus the complete experimental results can be summarized in terms of mean score coupled with a statement of the "order" of the binomial model.

1.5 Later, System M experimenters who had previously favoured the method of § 1.2 modified their analysis procedure. Like the method of § 1.4, the new method [Lessman, 1972] is based on a generalized distribution model linked to mean score. Unlike the method of § 1.4, the distribution model is based on the concept of a normal distribution, quantized according to the number of grades. This new analytical procedure is still being studied and evidence about its relative advantages and disadvantages has not yet been published.

A somewhat similar model has also been described [Allnatt, 1973] by a System I experimenter who employs the method of § 1.4. The model employs quantizing intervals which, although symmetrically arranged, are not equal. Calculations are greatly simplified by using the logistic function to calculate both the distribution curve and the quantizing interval. Analysis in terms of the model is comparable in complexity to the method of § 1.4. Match to the data is claimed to be better than for any other known all-embracing model, but the improvement compared with the binomial-type is slight. The benefit of the new model is most likely to be found in the many problems to be encountered in the application of results, where it will be advantageous to work in terms of distributions based on a notional continuous scale of opinion level.

1.6 In some cases it is only necessary to evaluate the perceptibility threshold of an impairment rather than the complete relationship between subjective evaluations and objective impairment. The "random stimuli" method is proposed [CCIR, 1974-78a]. Two slightly different definitions of the threshold may be used, either corresponding to the maximum standard deviation of assessments or to 50 per cent of the positive observations [CCIR, 1974-78b].

However, it should be remembered that knowledge of the threshold of perceptibility alone is insufficient in the treatment of any but the simplest situations.

1.7 Apart from the above fundamental differences, interpretation and comparison are further complicated by a number of points of detail:

1.7.1 The grading scale may be in terms of "quality" or "impairment", one being roughly, but in general not exactly, the inverse of the other. The number of grades in the scale has varied between 5 and 7.

1.7.2 Numerical scores are sometimes ordinal numbers which may either increase or decrease with the magnitude of the judged attribute. Alternatively the scores may be normalized so that the range of mean scores is from 0 to 1.

1.7.3 Objective impairment magnitudes may be expressed either directly or in such terms as signal/noise ratio. They may be in arithmetic or logarithmic (dB) units.

2. Subjective-impairment units

Although the present concern is with analysis of results, it is useful to take a note of a requirement that frequently arises in application.

Under practical viewing conditions, a number of impairments may arise simultaneously. As it is impracticable to test all the possible combinations, a "law of addition" of impairments can be of great benefit.

An empirical law, which has been used [Lewis and Allnatt, 1968; CCIR, 1970-74a], states that if $\bar{u}_1$, $\bar{u}_2$, ... $\bar{u}_r$ ..... $\bar{u}_n$ are the respective normalized mean scores for n unrelated impairments taken separately, the normalized mean score $\bar{u}$ for all impairments taken simultaneously is given by:

$$\frac{1}{\bar{u}} - 1 = \sum_{r=1}^{n} \left(\frac{1}{\bar{u}_r} - 1 \right) \qquad (1)$$

While a possible psychophysical basis for the law of addition in the above equation remains a matter for specualation /̲ Siocos, 1972 7̲ /̲ Allnatt, 1979 7̲, all the relevant experimental data available at the present time confirms that it appears to be valid to an accuracy sufficient for most, if not all, practical purposes /̲ CCIR, 1974-78c 7̲; although the multiple regression analysis quoted below may provide an interesting approach.

Consideration of design objectives is facilitated by expressing subjective impairment magnitudes in the form of directly summable quantities. As equation (1) suggests, this can be done very simply by transforming the mean score $\bar{u}$ into units of subjective impairment, termed "imps", by means of the relation $I = (1/\bar{u} - 1)$ imps. The I scale ranges from zero for the "perfect" picture ($\bar{u} = 1$) to infinity at the other extreme ($\bar{u} = 0$). $I = 1$ imp at the "mid-opinion" point given by $\bar{u} = 0.5$.

In the presentation of results, convenient mark points may be placed on the I scale, for example, 1/8, ¼, ½ and 1 imp. When the term "imp" was originally introduced [Lewis and Allnatt, 1965 and 1968], its proponents intended that it should be related to a particular quality grading scale and narrowly defined test arrangements [Prosser *et al.*, 1964; Corbett, 1970], with a view to providing as near an absolute basis as possible for results. Subsequently it has become clear that some qualification of results will almost always be necessary. For example, scales and observers' standards could vary from place to place, translation of the scale into another language produces an unknown effect, and sometimes it may be desired to apply the test technique to a television system for which broadcast-viewing conditions are unsuitable. It is suggested that details should be given of test arrangements which might affect results.

The following remarks offer some guidance to interpretation of the mark points when a particular set of conditions [Prosser *et al.*, 1964; Corbett, 1970; Allnatt and Bragg, 1968] is used. The 1/8 imp mark point represents a low level of impairment, of the same order as the residual impairment normally found with a laboratory set-up consisting of a high-grade slide scanner and picture monitor. Experience suggests that 0.25 imp may be taken as a practical design objective for each of the major impairments that may occur in a system, such as a national one, of moderate size and complexity. At the present time, 0.5 imp appears to be a reasonable design objective for each major impairment in a complex system involving transmission over a chain of long-distance international links.

One final work of caution is that the use of the "law of addition" to find the overall impairment resulting from the simultaneous presence of component impairments, should not be used to add impairments having the same subjective effect, but resulting from the errors of different objective parameters (e.g. in the PAL system, chrominance phase and chrominance gain errors both give rise to saturation impairments in the image). In this case, the overall impairment should first be calculated in objective terms from each of the errors of the individual objective parameters /̲ CCIR, 1970-74a 7̲.

A preliminary study of an alternative approach to the problem of multiple simultaneous impairments, has been reported [CCIR, 1970 74b]. In this, multiple linear regression analysis is shown to give satisfactory results with a set of available data. However, studies are continuing to examine the range of validity of the expression obtained, to examine the use of non-linear regression models and to extend the range of data used for checking the validity of the method.

Some new results /̲ Sallio, Kretz, 1982 7̲/, /̲ CCIR, 1978-82a, b 7̲/ suggest that the question of the applicability of Imp units to subjective cumulation is still open.

Further contributions to the study of the evaluation of the effects of multiple simultaneous impairments are invited.

3. Opinion distribution

3.1 *Statistical models of opinion distribution*

In §§ 1.4 and 1.5 reference was made to models based on the binomial and normal distributions. [CCIR, 1970-74c] summarizes the binomial type models referred to in § 1.4.; [CCIR, 1970-74d] summarizes, in its Annex, the data analysis method based on the normal distribution, and referred to in § 1.5.

It has been shown that the above two methods, to a good degree of approximation, can be directly related to one another as the logarithm of the impairment unit is both linearly related to the objective impairment and very nearly normally distributed, when the opinion distribution is taken as binomial [Siocos, 1972]. While this virtual correspondence should make it possible, if desired, to retain both the above methods of representing distribution, the advisability of recommending a single method should be studied.

3.2 *Double-stratified* technique

It is an important matter to examine the standard deviation of mean scores. Usually the test pictures need to be systematically selected, but sometimes it may be desired to use pictures drawn at random from a large population. In such a case, consideration should be given to the statistical properties of the double-stratified sample that may be obtained [CCIR, 1974-78d; Miceli and Orlando, 1977]. By this means, standard deviations can be evaluated and the results could suggest a possible economic compromise in future similar experiments, by proper choice of the relative numbers of observers and pictures used.

4. Use of comparison scale

A contribution [CCIR, 1970-74e] on comparison scales has shown that when the recommended 7-grade comparison scale is used for comparing pictures impaired by random noise, a more linear relationship between comparison grade and signal-to-noise ratio is obtained if the numerical weighting attached to the comparison grades is 4, 2, 1, 0, −1, −2, −4 rather than 3, 2, 1, 0, −1, −2, −3.

5. Prediction of the error in a subjective assessment test

If we consider the true subjective grade of a test condition to be the average obtained with a very large number of observations, it is possible to calculate the relationship between this true grade and the mean grade obtained by a practical experiment (with a realistic number of observations).

We can calculate the "expected error" associated with our number of observations at a defined confidence level. A, say, 95 % confidence level is a reasonable reference, and with this figure in mind it is possible to delineate the relationship between the number of observation and the maximum expected error (making certain statistical assumptions) /̲ Miceli and Orlando, 1977 7̲/ /̲ CCIR, 1978-82c 7̲/. This varies with mean grade and is given (for a mean grade of 3) in Fig. 1. Fig. 2 shows the correction factor to be applied for mean scores above and below 3.

The "expected error" can be visualized as the half-amplitude of the area of uncertainty within which the result of the test is expected to approximate the true value to be measured at the selected confidence level. From Fig. 1 we note that roughly equal numbers of pictures and observers give the lowest expected error for a given number of observations (see for example, points P and Q).

As an example of the application of the curves, if for instance a subjective test has been made with 9 observers and 8 pictures the maximum expected error is 0.47 (from Fig. 1, point Q). If the mean grade was 4.40, the correction factor (from Fig. 2, point Z) would be 0.75, so the "expected error" at 95 % probability would be 0.47 x 0.75 = 0.35. The curves were derived from analogue colour television experiments based on Rec. 500-1 and using reasonably critical pictures. They may however serve as a good approximation for other test conditions. It would certainly be helpful if some experiments were carried out to derive exact curves for the other cases.

The general principles given above always apply generally, but the degree to which the errors are in practice below the maximum, when critical pictures are used, is the subject of discussion. Some experiments point to successful application in this case, but other workers believe that the subject should be studied further.

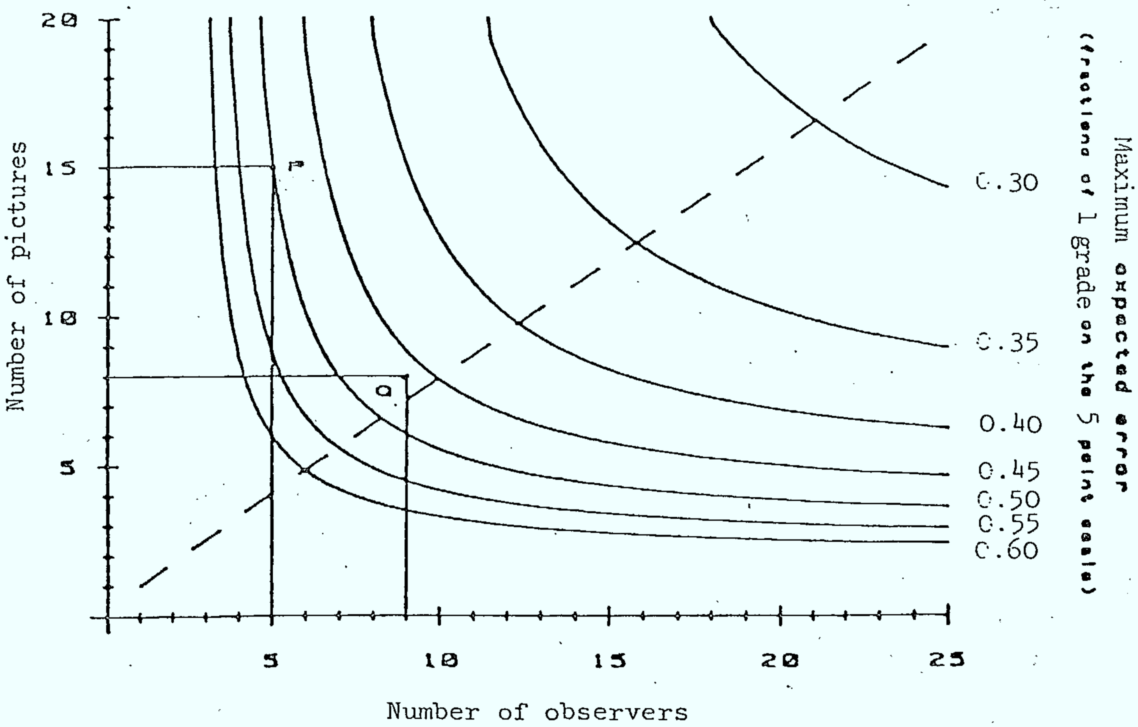

FIGURE 1

Curves giving the maximum expected error in a subjective assessment test, as a function of the number of observers and the number of pictures, for a confidence level of 95 %

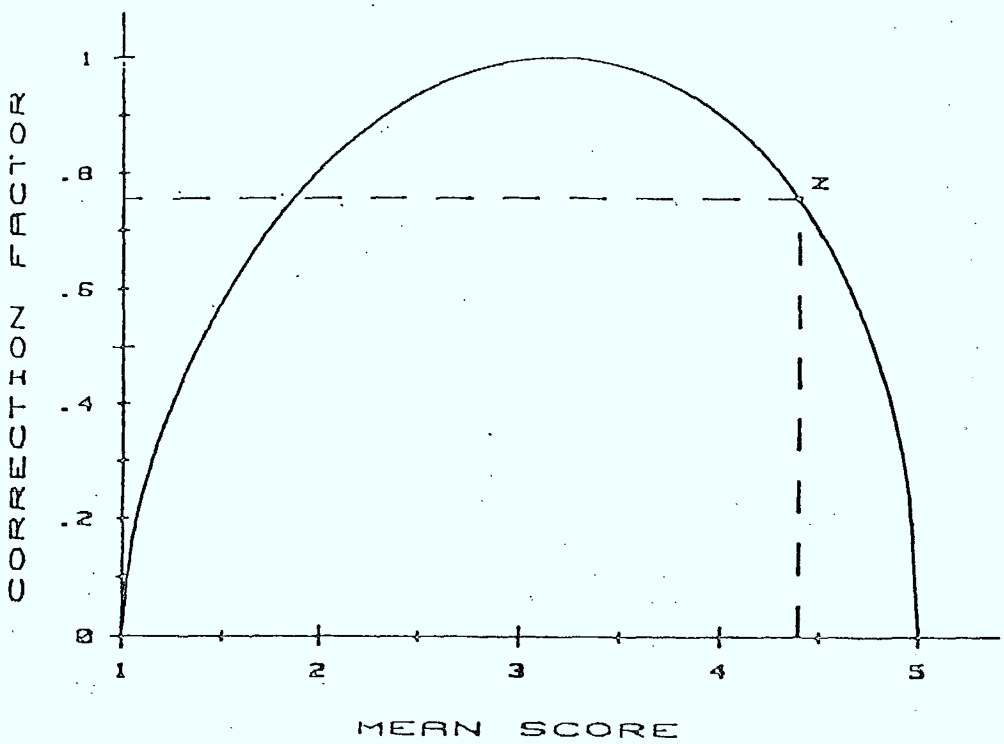

FIGURE 2

Curve giving the correction factor to be applied to the maximum expected error computed by means of Fig. 1, if the mean score is not 3

REFERENCES

ALLNATT, J. [1973] Opinion-distribution model for subjective rating studies. *Int. Journ. Man-machine studies*, 5, 1-15.

ALLNATT, J. W. and BRAGG, E. J. W. [1968] Subjective quality of television pictures impaired by sinewave noise and low-frequency random noise. *Proc. IEE*, 115, 317-375.

CAVANAUGH, J. R. and LESSMAN, A. M. [1971] Subjective effects of differential gain and differential phase distortions in NTSC colour television pictures. *Journ. Soc. Motion Picture Televis. Engrs.* 80, 614-9.

CORBETT, J. M. [1970] Effect of observer adaptation on the results of television quality-grading tests. *Proc. IEE*, 117, 512-514.

LESSMAN, A. M. [1972] The subjective effects of echoes in 525-line monochrome and NTSC colour television and resulting echo time weighting. *Journ. Soc. Motion Picture Televis. Engrs.*, 81, 907-916. Also, CCIR [1970-1974] 11/280 (USA).

LEWIS, N. W. and ALLNATT, J. W. [1965] Subjective quality of television pictures with multiple impairments. *Electronics Letters*, 1, 187-188.

LEWIS, N. W. and ALLNATT, J. W. [1968] Subjective and objective impairments in television pictures. *IEE Conference Publication*, 46, 2.1.1.-5.

MICELI, S. and ORLANDO, A. [1977], Subjective assessments elaboration for determining the distortion tolerance of a telecommunication system. Proc. of IEEE of EUROCON' 77, Venice, Italy, 1, 305.

PROSSER, R. D., ALLNATT, J. W. and LEWIS, N. W. [1964] Quality grading of impaired television pictures. *Proc. IEE*, 111, 491-507.

SIOCOS, C. A. [1972] Subjective impairment of television pictures — considerations regarding the summability of impairment units. *Journal of the SMPTE*, 81, 917-920.

SALLIO, P. and KRETZ, F. /¯April, 1982_7 to appear in EBU Review-technical. Comparison between two methods for the subjective assessment of television picture quality. Presentation of results with a single unit.

CCIR Documents

[1966-69]: a. CMTT/159 (Italy).

[1970-74]: a. 11/330 (Italy); b. 11/329 (Italy); c. 11/273 (United Kingdom); d. 11/279 (United States of America); e. 11/304 (Switzerland).

[1974-78]: a. 11/171 (Italy); b. 11/65 (France); c. 11/56 (United Kingdom); d. 11/421 (Italy).

/¯1978-82_7 : a. 11/313 (France), b. 11/331 (German Democratic Republic), c. 11/309 (Italy).

ANNEX III

EXAMPLE OF A METHOD OF PROCESSING THE RESULTS OF SUBJECTIVE TESTS

To provide a unified method of presenting the results of subjective tests using a particular cause of impairment, the following example describes the sequence of operations which can be carried out.

1. Processing test results for a particular value of impairment

The objective magnitude of the impairment is characterized by a particular number according to the convention appropriate to the particular effect; for example, signal/noise ratio in decibels, differential phase distortion in degrees, etc. This numerical value is called D and is defined more precisely in § 2.

The tests carried out with D constant, according to the principles set out in Recommendation 500-1, lead to a certain distribution of scores on the 5-point scale which is used. This distribution includes the differences in judgment between observers and the effect of a variety of conditions associated with the experiment, for example the use of several pictures.

In what follows U denotes the quality grade (1 to 5) on the scale being used.

The analysis of the distribution gives, for each grade U_i, the proportion of the total opinions p_i (note that $\sum_i p_i = 1$).

The analysis of this distribution can be made by two methods:

1.1 By simply calculating the direct mean score,

$$\bar{U} = \sum_{i=1}^{5} i \times p_i \qquad (2)$$

1.2 By finding a mathematical model (or interpolating function) which smooths out the random errors in the measured proportions p_i.

The different solutions to this problem which have been proposed fall under two main classifications. Firstly, models which yield directly the proportions in the five grades. An example of such a model is:

- a modified binomial distribution [Prosser *et al.*, 1964] which is described by its mean score and order (an integer).

The second class of models comprises those which are based on the concept of a smooth distribution function in terms of a continuous variable. Such a model can yield a distribution in terms of opinion grades, which can be matched to the experimental data, by a suitable process of quantization. Examples of such models are:

- the approximation by a Gaussian function, an example of which is given in [Lessman, 1972]. Such a function can be conveniently described by its mean and its standard deviation.
- the approximation by a logistic function [Allnatt, 1973]. The distribution is described by its median value and its width parameter (both being continuous variables).

Note 1. – The choice of method will depend both on the particular features of the distribution encountered in the experimental results, and also, on the mathematical treatment envisaged for the application of results.

Note 2. – In fitting the chosen model to the results for a given impairment condition, an efficient method of adjustment is simply to equate the mean and standard deviation of grades yielded by the model, to those actually found.

Note 3. – In an experiment with a given type of impairment, it is commonly found that a single value of width parameter (or binomial order, etc.) suffices to fit a well-chosen model to all the impairment conditions tested. Differences between the standard deviation actually found for a given impairment condition and that yielded by the model are no greater than can be accounted for by sampling error. Indeed, it has been found that a single width parameter can apply to many different types of impairment. Thus the finding of a distribution model with a width parameter which will apply to all impairment conditions in an experiment, and the establishment of a relationship between central tendency (mean or median) and objective magnitude can be treated as separate operations.

According to circumstances it may be appropriate for the next stage of the analysis to be made in terms either of the model mean (or median) or of the mean of the categorized distribution yielded by the model. It will suffice for what follows to have obtained the deduced mean score U_m, which may be identical to the direct mean score $\overline{U}$, but sometimes may differ slightly from it.

2. Processing to find the relationship between U_m and D

The next step in processing the data is to establish the relationship between the mean score U_m (or $\overline{U}$) and the objective measure of the distortion D as D is varied. The process which follows consists of finding a simple continuous relationship between U_m and D. The approximation of this experimental relationship by a logistic function is particularly interesting.

The processing of the data U_m can be made as follows:

2.1 The scale of values U_m is normalized by taking a continuous variable u such that,

$$u = (U_m - 1)/4 \tag{3}$$

when U_m is in the range 1 to 5.

2.2 Graphical representation of the relationship between u and D shows that the curve tends to be a skew-symmetrical sigmoid shape provided that the natural limits to the values of D extend far enough from the region in which u varies rapidly.

In other cases, where a physical limit on the value of D is close to the region in which u varies rapidly, a consequent marked dissymmetry of the curve can be observed.

In the latter cases, the variable D can be replaced by a function of D, such as the logarithm, which effectively moves the limit to infinity, thus achieving skew-symmetry of the curve. In what follows D is the value of the impairment expressed in units, usually logarithmic, chosen to yield skew-symmetry.

The use of arithmetic units, denoted by d, is common in the literature. The effect of this is discussed in the Appendix.

2.3 The function $u = f(D)$ can now be approximated by a judiciously chosen logistic function, as given by the general relation:

$$u = 1/[1/u_0 + \exp(D - D_M)G] \tag{4}$$

where G may be positive or negative.

In this expression the presence of the constant u_0* will be noted. It represents, in the function $u = f(D)$, the asymptotic value of u for the very small impairment.

This peculiarity of the curve $u = f(D)$ is frequently encountered and characterizes the fact that even for a level of impairment which is theoretically zero, the mean quality grade obtained in the course of the tests does not reach the limiting value $U_m = 5$, or $u = 1$. As a result of various imperfections of the picture, the measured value is u_0.

2.4 The value u obtained from the optimum logistic function approximation is used to provide a deduced numerical value I according to the relation:

$$I = (1/u) - 1 \tag{5}$$

In the case where the value u_0 is not equal to 1, this transformation provides a limiting values of I, as

$$I_0 = (1/u_0) - 1 \tag{6}$$

* In the publications this constant is also called H.

2.5 If I_{exp} is the raw result obtained from equations (4) and (5), a correction is made to this result by the relation, $I_u = I_{exp} - I_0$.

It is convenient in practice, for the graphical representation of the values of I_u as a function of the impairment D, to use a logarithmic scale for I_u. This convenience results from the fact that the logistic function being used to represent the relationship takes the form of a straight line in this graph. The expression for this relationship is effectively:

$$I_u = \exp{(D - D_M)G} \tag{7}$$

It will be observed that in such a graph the experimental values, processed to obtain corresponding values of I_u, will be sensibly on a straight line.

Interpolation by a straight line is simple and in some cases of an accuracy which is sufficient for the straight line to be considered as representing the impairment due to the effect measured by D.

Fig. 3 shows the representation thus obtained. The values of I are expressed, in accordance with the proposals which have been made by Lewis and Allnatt [Lewis and Allnatt, 1965], in "imp" units.

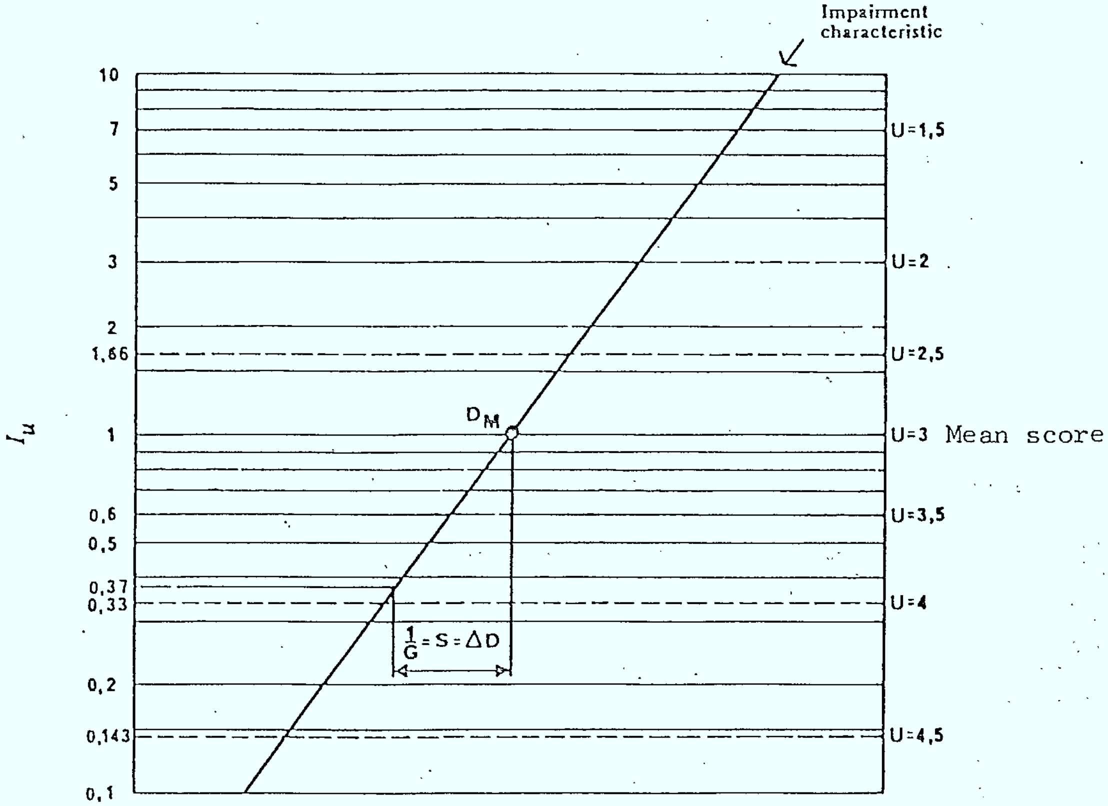

FIGURE 3 - Representation of the impairment characteristic
U - Mean score, 5-grade scale

The straight line may be termed the impairment characteristic associated with the particular impairment being considered. It will be noted that the straight line can be defined by the characteristic values D_M and G of the logistic function.

2.6 Adjustment for the residual impairments : an alternative approach

The relation $I_{exp} = I_u - I_o$ described at the beginning of § 2.5 is used to adjust the results for the residual impairment corresponding to u_o or I_o. In a comprehensive study / CCIR, 1978-82a, b, c /, / Kretz, 1981 / / Sallio, Kretz, 1982 /, comparisons between an absolute rather procedure and a procedure using direct anchoring with a reference unimpaired picture (see Annex IV, §§ 1.1.1, 1.2 and 3.1) provide another basis for the adjustment of the residual impairment. A shift applied to u values ($u' = u + 1 - u_o$) is found to correspond very satisfactorily to the data. According to this analysis equation (4) would become :

$$u = (u_o - 1) + 1 / [1 + \exp (D - D_M) G] \qquad (4')$$

then, an Imp value can be calculated from adjusted u' data, by setting :

$$I_u = (1/u') - 1 \qquad (5')$$

and equation (7) is seen to be also applicable, the rest of the analysis remaining unaltered.

However, it should be noted that this technique fails when scores are very low, and thus cannot be regarded as satisfactory for all purposes.

3. Conclusions

The processing of subjective test data which is described above has two advantages:

- the impairment characteristic is a straight line;
- in very many cases, the values of *I* appear to possess the property of summability when several sources of impairment which are not correlated are present simultaneously on the picture.

It is desirable that for each of the causes of impairment known to be important, the results of a subjective test should be available as an impairment characteristic presented in this form. The establishment of a collection of these characteristics would facilitate the use of the results of subjective tests.

APPENDIX

ARITHMETIC MEASURE OF OBJECTIVE IMPAIRMENT

Where the conventional objective measure of impairment is arithmetic in nature, the relationship with u will generally not be skew-symmetrical and it will be necessary to employ logarithms, as discussed in § 2.2 of Annex III. However, the arithmetic measure, which will be designated d, might be considered the more basic one and it is of interest to examine the function relating u with d.

$$D = \log_e d$$

and,

$$D_M = \log_e d_M$$

Substituting d in the logistic equation (1), it can be shown that,

$$u = 1/[(1/u_0) + (d/d_M)^G] \qquad (8)$$

from which (as in § 2.5 of Annex III),

$$I_u = (d/d_M)^G$$

Although the relationship of u to d is not skew-symmetrical it will be seen to be extremely simple in form.

It may be considered that the corresponding values of G (or equally those arising from values D which are natural logarithms) are of more fundamental significance than those from values of D expressed in decibels, or a logarithmic measure other than natural logarithms. Accordingly, it is proposed to reserve G exclusively for values appropriate to equation (8), as this usage is common in the literature and particularly because an extensive set of results, expressed in these terms, has been published, [Report (AC/11)].

Values of G arising from values of D expressed in decibels may be conveniently designated G' and are related to G as follows:

$$G = (20 \log_{10} e)\, G' \qquad (9)$$
$$= 8.686\, G'$$

REFERENCES

ALLNATT, J. [1973] Opinion-distribution model for subjective rating studies. *Int. Journ. Man-machine studies*, 5, 1-15.

LESSMAN, A. M. [1972] The subjective effects of echoes in 525-line monochrome and NTSC colour television and resulting echo time weighting. *Journ. Soc. Motion Picture Televis. Engrs.*, 81, 907-916. Also, CCIR [1970-1974] 11/280 (USA).

LEWIS, N. W. and ALLNATT, J. W. [1965] Subjective quality of television pictures with multiple impairments. *Electronics Letters*, 1, 187-188.

LEWIS, N. W. and ALLNATT, J. W. [1968] Subjective and objective impairments in television pictures. *IEE Conference Publication*, 46, 2.1.1.-5.

PROSSER, R.-D., ALLNATT, J. W. and LEWIS, N. W. [1964] Quality grading of impaired television pictures. *Proc. IEE*, 111, 491-502.

KRETZ, F. [September-October, 1981] Revue Radiodiffusion-Télévision, Représentation des resultats d'essais subjectifs par une unité unique.

SALLIO, P. and KRETZ, F. [April, 1982] to appear in EBU Review-Technical. Comparison between two methods for the subjective assessment of television picture quality. Presentation of results with a single unit.

WHITE, T.A. [1980], Transmission of alphanumerics by television : assessment of typescript by "experts", Proc. 9th International Symposium on Human Factors in Telecommunications, New Jersey, 27-34.

WHITE, T.A. [1981], Transmission of alphanumerics by television. Displays, 2, 295-299.

ALLNATT, J.W. [1979] Television measurements through psychophysics to subjective picture quality, Radio & Electronic Engineer, 49, 611-619.

CCIR Documents :

[1978-1982] : a. 11/71 (France), b. 11/257 (France), c. 11/312 (France).

ANNEX IV

TEST PROCEDURES

This Annex describes various test procedures for organizing subjective assessment tests. Recommendation 500-1 gives much information about subjective assessment but the procedure still has parameters which are open. § 1 of this Annex describes different procedures in some detail. § 2 describes the virtual sample technique. § 3 gives information on results obtained by some of the procedures in comparable situations. § 4 gives some additional evidence. § 5 discusses different arguments related to the choice of an approximate method for various situations.

1. Description of procedures

This section presents descriptions of various procedures. The descriptions cover such items as grading scales, design of the presentation sequence, definition of an impaired condition and so on. The validity of these items is not necessarily restricted to the particular procedure being described.

1.1 Procedures using quality rating scale

1.1.1 Single-stimulus procedure using quality grading scale

This procedure is essentially that of /̅ Prosser et. al., 1964_7. It accords with Recommendation 500-1 and employs the five-grade quality scale. It has been in continuous and satisfactory use by British Telecom since 1962, and procedures which were substantially identical have been successfully applied by certain other organizations in several countries.

It is primarily intended for the assessment of impairments over a full range of subjective magnitudes. Although it can be employed with smaller ranges, it is known /̅ Corbett, 1970_7 that the rating standards employed by observers may then be different. One possibility of dealing with restricted-magnitude impairments which has been shown to be viable /̅ Kitchen, 1977_7 is to 'embed' the test impairment(s) among a number of other impairment conditions, similar in character but with the preferred full range of magnitudes.

Tests are conducted in a specially equipped viewing room, three sides of which are draped in gathered white curtains, with the fourth side (to the rear of the observers) draped in grey curtains to prevent specular reflections. The lighting provides a colour temperature of 6500 K. Care is taken to ensure that the performance of the basic test channel is of the best possible quality. Display brightness and contrast are standardized using a 'PLUGE' test signal*. Use of PLUGE leads to a display conforming with Recommendation 500-1, but provides more complete standardization.

---

* PLUGE (Picture Line-Up Generator, Electronic) is a device designed by the BBC /̅Anstey and Ward, 1962_7. For the current purpose, the unit is set so that, with the video output signal at the standard peak-to-peak amplitude of 1V, the white level is 0.7V above blanking, and the vertical bars are at ±0.02V relative to 'black' level, which in turn is 0.02V above blanking level.

Generally, twenty or more observers are employed, non-expert in the sense that none work in television engineering, or in the photographic or allied fields involving visual arts. Immediately prior to a session, the observers are screened for visual acuity using appropriately scaled Snellen charts, and (except for monochrome tests) for colour vision using Ishihara charts /_ Ishihara, 1949_/.

At the start of each session, recorded instructions are given, observers being asked to rate the quality of each picture condition in turn, using the following 5-grade quality scale :

A ... (Excellent)

B ... (Good)

C ... (Fair)

D ... (Poor)

E ... (Bad)

Generally the scale has been used to rate overall picture quality, but it can also be used for other characteristics, e.g. the quality of legibility /_ White, 1980; White, 1981_/.

No numbers are associated with the grades on the record sheets used by the observers. The scale is illustrated, in the form shown above, at the top of each sheet. Observers record each opinion by ticking the appropriate square of a matrix, in which the rows are labelled A to E, and the columns correspond to the successive picture-conditions that are displayed, termed 'presentations'. Two such matrices, catering for a total of 40 presentations, are conveniently contained on one side of an A4 sheet together with the grading scale.

A number of demonstration pictures containing impairments typifying those to be assessed are first presented. They can, if desired, be selected from among the picture conditions in the test proper. It is pointed out that these preliminary demonstrations typify what is to be seen in the test itself but do not necessarily cover the full range of conditions. Observers rate the demonstrations but these assessments are not included in the main analysis.

Where the experiment is of a suitable type, the observers are offered as long as they wish to form their opinions. Normally 10 s is found sufficient for the display of each picture condition, irrespective of whether the pictures are stills or moving sequences. Pictures are removed for about 15 s between successive presentations, while observers record their opinions. Thus the experiment may proceed at a uniform rate. If there are more picture conditions than can be dealt with at a single session lasting about half an hour, observers may be recalled on subsequent days for up to three further sessions. It is often preferable to employ completely fresh observers, if these are conveniently available.

So far as possible, the tests are based on a randomized-block design, derived from Graeco-Latin squares /_ Fisher and Yates, 1970_/ to balance the effects of all factors evenly over the experiment as a whole. Where the desired structure does not permit a complete balance, an important requirement is that it should be possible to assess any adaptation effects in the observers that occur during both the preliminary demonstrations and the main part of the experiment.

The test conditions are selected to give an estimated range of quality and a grand mean score as specified in § 2.5 of Recommendation 500-1. If the grand mean value of the original scores actually achieved falls outside the range 2.92 to 3.08, consideration should be given to repeating the experiment with suitably modified impairment levels.

1.1.2 Double stimulus procedure using pairs of continuous quality-rating scales

This procedure was announced in /¯White and Allnatt, 1980_7. It is intended for use in tests with a limited range of impairment magnitudes, and in particular where the impairments are very small, as in digital codecs of good quality. It avoids the effort and expense of obtaining results having no particular application, when the 'embedding technique' (see § 1.1.1) is used with the single-stimulus quality-grading procedure. The procedure is still in an experimental stage, but preliminary tests by British Telecom and the United Kingdom Independent Broadcasting Authority have provided encouraging results.

The procedure invokes the concept of comparison with an unimpaired picture, but defers explicit comparison until the analysis stage. One condition in every pair is always the unimpaired reference picture, but its position is changed in quasi-random fashion and subjects are not told which of the two it is. The observers rate each picture individually on an absolute scale. Although they are able to make comparisons between each pair of pictures if they wish, they are not asked to do so. It is because the picture conditions are assessed in pairs, that the procedure is termed a double-stimulus procedure.

As any improvement in precision due to the availability of a reference condition would be reduced by quantizing error if any of the grading scales of Recommendation 500-1 were to be used, a continuous quality-rating scale is used. The scale has been arranged to resemble the existing quality-grading scale (Table I of Recommendation 500-1) as closely as possible. The score sheet contains pairs of such scales each 100 mm high, spaced 10 mm apart, the space between the members of each pair being 5 mm. The terms categorizing the intervals appeared only on the left of the first scale. For more information see /¯CCIR, 1978-82e_7. Observers register their opinions by marking the scales in order, corresponding to successive picture-conditions.

The desiderata, expressed in § 2.5 of Recommendation 500-1, that the grand mean score should be close to the value at mid scale does not apply. However, the specifications in Recommendation 500-1, supplemented by those in § 1.1.1, are applied as far as possible.

In the initial experiments by British Telecom using the method /¯White and Allnatt, 1980; Redstall and White, to be published_7, each observer attended separately, and was allowed to switch instantaneously between the two picture conditions until he was satisfied that he had established his opinion of each. In the UKIBA experiments /¯CCIR 1978-82f; IBA, 1981_7, the observers attended in small groups, and the display was instantaneously switched a number of times between the two conditions according to a pre-arranged plan.

1.1.3 Procedure using the quality scale with direct anchoring*

To study the role of a reference picture, a new method has been designed /¯CCIR 1978-82c_7, /¯Kretz and Sallio, 1981_7. In this method, the 5-grade quality scale from Recommendation 500-1 is used and the procedure is the same as for the procedure of the EBU method (see below § 1.2) : a reference picture is displayed before each picture to be assessed and the observers are told to rate the pictures in relation to the reference which is said to correspond to grade 5 ("excellent"). With this procedure, direct anchoring of the rates at the top of the scale is obtained with the quality grading scale.

1.1.4 Procedure using the quality scale with high and low indirect anchoring

This method /¯CCIR, 1978-80o_7 is intended to give indirect anchoring at both ends of the scale. It uses a number of unimpaired pictures, and a number of heavily degraded pictures which are presented and judged during the test in the same way as the other pictures. The anchoring effect is only indirect because the observers are not informed of the presence of the top and bottom reference pictures.

Indirect anchoring with a lower quality reference could be applied to other procedures described in this Annex.

As the indirect anchoring is applied to both ends of the scale, full range impairments need not necessarily be used in a given session, and it is expected that there will be no need to standardize the grand mean score as required by Recommendation 500-1. A complete experiment should be arranged as follows :

- in order to keep the time required for sessions as short as possible, not more than 2 or 3 values of the distortion to be investigated should be tested in the same session. So several sessions are usually needed for the complete experiment and for each session it will be better to use a new group of observers.
- for each value of the distortion a group of 5 or more pictures should be chosen.
- two groups of 4 or more pictures each, to serve respectively as high quality and low quality references, should be chosen as well.
- the impairment used for the low quality reference must correspond to a score of approximately 1.5. The addition of white noise at the level of 19 dB S/N ratio could be used for the purpose.
- the usual care should be taken during the presentation (random sequences, etc.).
- /¯CCIR, 1978-82l_7 suggests avoiding repeated use of the same picture several times in the course of the presentation, in order to keep biasing as low as possible. This suggested approach can be applied to any of the procedures described in this Annex.

---

* Some for of anchoring is always implicit in all the procedures in the sense that it is necessary to standardize the rating process. Here the term "direct anchoring" refers to explicit anchoring. "Indirect" corresponds to standardizing the adaptation phenomena /¯Corbett, 1970_7 by means of the range of impairements in a test session.

1.2 Procedure using the impairment scale and a reference picture (EBU method)

The EBU requiring directly comparable results from different laboratories on subjective quality assessments has agreed a method using the 5-grade impairment scale (Table I of the Recommendation 500-1) and a reference picture. This procedure conforms to Recommendation 500-1 but specifies the test conditions in more detail. It was designed mainly for the assessment of digital television pictures but is not restricted to this application. The use of a reference is to give better precision /‾ i.e. standard deviation_7 for low impairment situations and the use of the impairment scale is to obtain a direct measurement of the perceptibility threshold as well as larger impairment magnitudes.

The method is described in more detail in /‾ CCIR, 1978-82a_7 and /‾ Bernath et al., (1981)_7. The general principles were drawn from /‾ Sallio, Kretz (1978)_7 and /‾ CCIR, 1974-78_7. It has been used successfully by various EBU members as well as recently by the SMPTE (Bennett, 1981) and Japan /‾ CCIR, 1978-82n_7.

1.2.1 Conditions

The reference picture used in the procedure is an unprocessed picture formed (for 625-line systems) from either R, G, B (each 5.8 MHz bandwidth) or Y, U, V (5.8 MHz bandwidth for Y and 1.5 MHz bandwidth for U and V) components. Once only coded PAL or SECAM can also be used, when necessary, but the same reference standard should be used throughout each session.

When only still pictures are used, the EBU has selected nine 35 mm slides*, primarily intended for digital studies, from which most of the test material is drawn (no set of test moving sequences has yet been selected). The source is adjusted for the best picture quality corresponding to the standard defined for the reference picture; the technical characteristics of the equipment used, and the adjustment made are included in the statement of the results. Viewing distances of both four and six times picture height are used. Viewing is on the best available delta or in-line colour monitor of large size (at least 50 cm diagonal) and details and adjustments made on the monitor are stated in the results (for more information on this matter see Annex VII).

1.2.2 Observers and instructions

Observers (non-experts who have normal vision) are introduced beforehand to the types impairments on still pictures not used for the real experiments, and to the grading system, but without trying to explain the meaning of its wording. Identical neutral-voiced pre-recorded instructions are given at the beginning of each viewing session, and every precaution should be taken to avoid distracting the observers by the visual and auditory environment, during the session. The first session is a preliminary "warm-up" session, and the results obtained at this session are later ignored.

---

* Now available from EBU.

1.2.3 Procedure

Each session corresponds to a sequence of presentations. A presentation uses the same picture, and has four phrases : the reference picture period of 10 seconds, a mid-grey separation period of 3 seconds, the test condition period (10-15 seconds, depending on the complexity of the impairment to be assessed) and a mid-grey period for 10 seconds for voting. Voting is limited to the last period. Each test presentation is given twice in each test session to check the coherence of each observer. The test pictures and impairment conditions are in a random order which changes from session to session, but avoid two successive presentations with the same visual scene. In a session, the reference picture is assessed twice, as a test condition.

The total session does not last more than 30 minutes, and the same observers have at least a half-day gap between two sessions. The two votes obtained on the same test condition and picture by the same observer which differ by more than two grades are omitted. If more than 15 % of the assessments for a test condition are so omitted, then it is stated (but generally it is not more than 5 %). Results are presented separately for each viewing distance, picture, etc., unless quite unnecessary. For each test parameter, the mean score and the standard deviation of the subjective assessment are stated.

2. The virtual sample technique

2.1 The main sources of errors in subjective texts are essentially two :

- random (i.e. stochastic) errors, and
- systematic errors.

Provided the objective conditions of the test have been standardized, the nature of the errors is purely related to the parameters used in the design of the tests (number of observers, pictures, procedure used, etc.).

Stochastic errors are very easy to recognize. In the case of a relationship between distortion magnitude and picture-quality, they lead to some dispersion of the experimental mean scores around the fitted curve (e.g. obtained by least spurious method). Standardization of test procedures usually tends to reduce stochastic errors but due to their nature they can be further reduced by statistical averaging (the use of a fitted mathematical function is one kind of averaging).

Systematic errors are difficult to recognize because they are almost unrelated to random errors. Usually they act by shifting the previously mentioned curve (biasing) and/or affecting its slope. Once a certain degree of systematic errors have been introduced in the experimental results, they cannot be averaged out by statistical methods.

The sources of systematic errors are essentially two :

- the observers and pictures used in the experiments;
- the way by which the experiments are carried out (procedures).

The effect of the former might be reduced by increasing the number of the observers and pictures, but the experiments may become impractical.

.1 Comparison of the results obtained by single stimulus method and EBU method

The two procedures described in § 1.1.1 and § 1.2 of this Annex have been subjected to a series of comparative experiments [CCIR, 1978-82b, c, p] [Sallio and Kretz, 1982]. Different types of impairment were used : analogue filtering, additive noise (in two different contexts), cross-modulation transfer noise, edge-busyness (in two different contexts), accumulation of additive noise and edge-busyness transmission errors at 34 Mbit/s (in two different impairment ranges). Twenty independent groups of ten non-expert observers (one group for each type of impairment and each procedure) participated in the tests (a total of 46 sessions was necessary for each method). The comparisons of the results obtained were analyzed in terms of average grades and standard deviation, separately for each viewing distance. The following conclusions were drawn :

- the two methods lead to objective-to-subjective relationship curves with very similar curve shapes. There is a shift between the characteristic obtained by each method. The quality rating curve tends to be rather below impairment ratings curve [EBU method]. At mid-scale (grade 3), the standard deviations of the scores are at maximum and are very close for both procedures (at 6 H, 0.84 for the single-stimulus method and 0.79 for the EBU method).

- the impairment scale procedure using a reference picture produces an average grade for reference picture very close to the highest grade (4.88 on average), showing good anchoring for the ratings at the top of the scale.

- the quality scale procedure leads to an average grade for reference pictures which varies from almost 0 to one grade below the highest grade (4.56 on average). This seems to be due to the absolute nature of the rating.

- the two methods tested are both sensitive to the context and the range of the impairments presented in a session. These two subjective phenomena have identical effects on the two methods. It would therefore seem important in presenting the results, to describe the precise conditions (range of impairments presented in each session, context within a session); this allows a better understanding of the results.

- in the range between "imperceptible" and "perceptible but not annoying" (around grade 4.5), the procedure using impairment scale and a reference, leads to standard deviations of the scores 1.4 times smaller than with the quality scale procedure; thus the former procedure seems to give better precision, and might allow the number of scores toward the top of the scales to be halved.

These results suggest that it may be possible to obtain a transformation of quality mean grades into impairment mean grades obtained using the EBU procedure by shifting the experimental results by the amount associated with the residual impairment (mean grade for unimpaired pictures). The transformation of impairment mean grades into quality mean grades is suggested by shifting the experimental values by half a grade, although this value is not completely stable (see also Annex III of this report). Due to the bounded nature of the scale, these transformations cannot apply at the bottom end of the scale.

The technique of the "virtual sample" has been primarily intended for the reduction of systematic errors. Such a technique might be thought to increase the amount of random errors, but, in the final analysis (i.e. a curve obtained by least squares fit) they will be averaged.

For more complete details see /¯Miceli and Orlando, 1977_7, /¯CCIR, 1978-821_7.

2.2 The "virtual sample" is constituted by a relatively large number of observers (e.g. 50 or more) and also by a large number of pictures (e.g. 40 or more). It is called the "virtual sample" because it is not used in its entirety for the actual assessment tests; rather, it is used as a population from which small samples of manageable size are repeatedly drawn.

Taking the example in which a relationship between given distortion magnitudes and picture quality is to be found, and having in mind that the aim of the virtual sample technique is to use different samples of observers and pictures for different test conditions, the complete planning of the experiments when following the cirtual sample technique should be as follows :

- a number of test conditions (e.g. 8 to 10 values of the distortion magnitude) should be selected;
- a number of groups each made up of no more than 2 or 3 non-contiguous test conditions should be formed;
- for each test condition, a random sample of 5 to 6 test pictures is chosen from the complete set (virtual sample), so each group of test conditions will have 2 or 3 such sets of test pictures. For each group a sample of 8 to 10 observers should be selected from the complete set (virtual sample). In this way we have different test pictures for different test conditions in the same group;
- for each group of test conditions, a session following the procedure given in in § 1.1.4 should be arranged;
- the mean scores should be calculated and then fitted by least squares using a suitable function (e.g. the "logistic function");
- some statistical tests should be carried out for the goodness of fit, and the final output of the experiment will be the fitted curve just obtained.

If more accuracy is required for the fitted curve, the experiment may be repeated as a new experiment, and the corresponding mean scores should be averaged with the past ones before performing the new least squares fit.

There is general agreement that the technique is appropriate for complex impairements at least in relation to observers. However, some administrations believe it may not be economical in the case of a single impairment.

3. <u>Results of directly comparable experiments</u>

This section describes results obtained with different assessment procedures applied to the same experimental material (pictures and impairments). Naturally, when a method is used in a specific situation, account must be taken of numerous parameters and the conclusions drawn from such experiments, described in this section, take only some of these factors into consideration. This is discussed in detail in § 5.

3.2 Comparisons of results obtained with other methods

To study in more detail the role of reference pictures, anchoring and grading scale, several methods have been tested on the same type of impairment / CCIR, 1978-82c /. This study consisted of a comparison of the results obtained with two previously compared methods (§ 3.1) and the results obtained with various other methods. The following aspects were investigated :

- the use of the 5-grade impairment scale and of a continuous impairment scale, of the 5-grade quality scale and of a continuous quality scale, all with a reference picture for direct anchoring (in effect the EBU procedure using different scales);

- the use of a continuous quality scale using a double stimulus procedure close to the one described in § 1.1.2.

The following conclusions were drawn :

- good anchoring is possible in the upper part of the range with an EBU type of presentation but using a 5-grade quality scale and informing the observers that the reference picture should correspond to the grade "excellent";

- the comparison of results obtained by methods which differ only in the use of a discrete or a continuous rating scale shows that neither the continuous quality scale nor the continuous impairment scale provides more information than the recommended 5-grade scales (comparable means and standard deviations);

- with double the stimulus method which does not provide direct anchoring, the mean score for reference is not close to the top of the score range. The standard deviations obtained by this method are not significantly lower than those obtained by the method recommended by the EBU;

- the implementation of continuous scales gives rise to some problems : certain observers (non-experts) find them difficult to use and, analyzing and presenting the results of the experiments, becomes more complicated.

3.3 Comparisons of the results obtained using 5- and 6-grade quality and 6-grade impairment scales

The comparison of the performance of the 5-grade quality scale and a 6-grade impairment scale have been reported in / Allnatt and Corbett, 1974 /, and recently re-examined with particular reference to performance near the threshold of visibility / Allnatt, 1980 /. The procedure was the same except for the scale used. The 6-grade impairment scale differs from the present Recommendation 500-1. Two types of impairment were considered, one with 625-line monochrome television using a 2 μs undistorted echo, and the other with opaque photographs impaired by blur. Results were analyzed only in term of mean score. The main conclusion of this study is that the impairment scale offers no advantage in respect of sensitivity with impairments below the threshold of visibility. However, it should be pointed out that this experiment does not represent results that would be obtained by comparing the single stimulus method and the EBU method.

4. Some other experimental evidence

Experimental evidence from other sources is available, some of it differing from that in § 3. A considerable body of results exists in the case of the single-stimulus quality-grading method, relating to all its important properties.

As regards the relative value of the quality scale and the impairment scale, some see the "quality" concept as more closely related to viewers interests, and, furthermore, beneficial because if an "impairment" actually improves the picture, the results reflect this. On the other hand, others see the impairment scale as easier to interpret and as having the advantage of allowing measurement of a threshold of perception (between grade 4 and 5 on the impairment scale). There is evidence of a balance of preference of observers opinion between the two scales. There is evidence that it may be possible, in some or all cases, to relate the two semantic axes by an appropriate formulation, and work is continuing in this field.

The arguments for a continuous scale are that sometimes the extra organization and analysis time are justified because fine discrimination is both possible and needed. The arguments for a discrete scale are that no better results can be achieved with a continuous scale, for reasons that include the fact that non-experts are, in practice, no more discriminating than the discrete scale allows.

The need for some form of anchoring is recognized by all, and this can be achieved in various ways. Some workers suggest that for impairments which are of interest over a wide range of values, no specific regular reference picture is needed, because the standardized wide range of impairments itself causes observers correctly to orientate themselves on the scale. For experiments using only very small impairments an anchor picture should be provided, but it should not be signified as such, because this would make the observation environment too artificial. The double stimulus procedure may have an advantage in measurements of impairments of small magnitude such as in future systems. Other workers argue that the regular, signified, use of a reference picture (high-quality) helps observers to orientate themselves on the scale and that the results of experiments demonstrate this, particularly for small impairments. Another approach which is believed to be valuable is to use two unsignified reference pictures (high and low-quality).

As regards the EBU method, a grade for the reference, in one case as low as 4.63 was found. In a work by the Australian Broadcasting Commission, a grade for the reference of 4.42 was noted in one case. In SMPTE assessments, using NTSC and RGB as references with observers taken from the broadcasting industry, reference picture grades of 4.7 were noted. Some workers believe that the reason for low values may include insufficient procedural control or non-consistent reference quality, as related to the necessary direct anchoring. In those cases, correction for residual impairment seems necessary as for single-stimulus procedurs. Further studies in relation to other procedures may be relevant.

In a move towards rationalization of methods, it should be possible, in the next CCIR Study Period, to bring major elements of the procedures together and confine alternatives to only certain parts. It is encouraging to note that in a recent practical situation (the study of the relationship between impairment and different digital sampling frequencies [CCIR, 1978-82g, h, n]), where careful procedural rules were applied, virtually the same mean scores were achieved by entirely independent tests using different procedures (§ 1.2 and § 1.1.2 above).

Results support the applicabilities of the Imp transformation (see for example Macdiarmid and Allnatt, 1978; also Annex II) as a law of addition of subjective impairments. This law is used to adjust for the effect of residual impairment in analysis, and operates differently from the transformation using shifting of the mean scores, advocated in § 3.1.

As regards the double-stimulus quality grading method, the standard deviations of differences between pairs of scores at the same presentation have been found to be lower than those of the individual scores, when impairments are small. A transformation is required to give the equivalent standard deviations in terms of a 5-grade scale, and at zero impairment values of about 0.13 were found with a fairly wide range of impairment using random noise (White and Allnatt, 1980), and 0.35 in assessment of a high quality digital codec (Pedstall and White, to be published). In other experiments with digital television (IBA, 1981) the value was 0.22. Comparable values found in similar experiments /̅ Kretz and Sallio, 1981_7 were 0.25 at 4 H, 0.45 at 6 H for zero impairment.

Comparison /̅ Redstall, White, 1982_7 of measurements of a digital codec using the double stimulus method with those using the single stimulus method with the test impairments embedded among a number of other wide-range impairments, support the finding of /̅ White and Allnatt, 1980_7 that adaptation due to indirect anchoring effects are considerably reduced by using double stimulus as compared to the single stimulus method.

5. Discussion

The aim of specifying a procedure in detail is to minimize the random variation of results that is not due to systematic differences between different populations of observers found when, say, the results of independent tests are compared or combined. This is the common aim of all the procedures described. But along with this are other factors which are seen as giving strength to particular procedures. These relate to the degree of discrimination of results that is worthwhile, and the degree of complexity and elaborateness of the test procedure. The more complex and elaborate the procedure, the more time consuming and costly it may he. How much is to be gained in terms of the results achieved and the accuracy needed ? These are central to the discussion on the choice of procedure.

The procedures described in § 1 select a combination of quality or impairment scale with the regular use of a reference picture, either signified as such or not, or indirect anchoring by impairment range. Procedure 1.1.1 is the simplest method for organization; the analysis of results for 1.1.1, 1.1.3 and 1.2 are about the same complexity; the organization of 1.1.2, 1.1.3 and 1.2 are about the same, but the analysis of results for 1.1.2 takes longer. Method 1.1.4 generally requires more sessions than the other procedures and is designed to reduce systematic errors. A comparison of results obtained by some of the methods for certain impairments is given in § 2, and other experimental evidence in § 4.

It is clear that each method is seen as having certain strengths and that the choice among them is not a simple one. It is impossible to do full justice to the arguments in a report of this kind, but essentially the main arguments which have influenced workers in the field are as follows.

The choice of procedure is related to the choice of grading scale, whether it should be continuous or discrete, and the way in which observers should be asked to make the correct use of the scale.

REFERENCES

ALLNATT, J.W., Electronics Letters, 1980, 16, 450-451, Subjective assessment method for television digital codecs.

ALLNATT, J.W. and CORBETT, J.M. PROC IEE, 1974, 121, pp 785-793, Comparisons of category scales employed for opinion rating.

ANSTEY, H.G., and WARD. P. [1963]: Operational control of television studio picture quality, IEE Conference Report Series No. 5, 286-292.

AUSTRALIAN BROADCASTING COMMISSION [1981] Tests of Subjective Impairment due to Random Noise (Report to be published).

BENNETT, D. [1981] SMPTE component-coded digital video picture quality assessments, SMPTE Jnl, October 1981, 960-967.

BERNATH, K., KRETZ, F., WOOD, D. [April 1981], The EBU method for organizing subjective tests of television picture quality, EBU Technical Review, No. 186.

CORBETT, J.M. [1970]: 'Effect of observer adaptation on the results of television quality-grading tests', Proc IEE, 1970, 117, pp 512-514.

FISHER, R.A. and YATES, F. [1970]: Statistical tables for biological, agricultural, and medical research, Oliver and Boyd, Edinburgh.

IBA [1981] Experimental and Development Report 114/81.

ISHIHARA, S. [1949]: Tests for colour blindness, H.K. Lewis, London.

KRETZ, F. and SALLIO, P. [September-October, 1981] Radiodiffusion-Télévision, Comparaison de deux méthodes d'évaluation subjective de la qualité des images : rôle des images de référence, de l'ancrage et de l'echelle de notation.

KRETZ, F. [September-October, 1981] Radiodiffusion-Télévision, Réprésentation des resultats d'essais subjectifs par une unité unique.

KRETZ, F. and SALLIO, P. [April, 1982] To appear in EBU Review-technical. Comparison between two methods for the subjective assessment of television picture quality. Presentation of results with a single unit.

MACDIARMID, I.F. and ALLNATT, J.W. [1978] Performance requirements for the transmission of the PAL coded signal. Proc. IEE, 125, 571-580.

MICHELI, [illegible], ORLANDO, A. [1977], Sampling procedures and goodness of Fit. Proc. of [illegible] on measurements, Lannion, France, Oct. 1977.

PROSSER, R.D., ALLNATT, J.W. and LEWIS, N.W.: [1964] Quality grading of impaired television pictures, Proc. IEE, 111, pp 491-502.

REDSTALL, M.W., and WHITE, T.A. /‾1982_7, Subjective quality of a 70 Mbit/s digital codec for colour television, (to be published).

SALLIO, P. and KRETZ, F. /‾April-May, 1978_7 Qualité subjective en télévision numérique. Première partie : méthodologie de son évaluation. Revue Radiodiffusion Télévision N.52, 13-19.

WHITE, T.A. and ALLNATT, J.W. /‾1980_7: 'Double-stimulus quality rating method for television digital codecs', Electronics Letters, 16, pp 714-716.

WHITE, T.A. /‾1980_7, Transmission of alphanumerics by television : assessment of typescript by "experts", Proc. 9th International Symposium on Human Factors in Telecommunications, New Jersey, 27-34.

WHITE, T.A. /‾1981_7, Transmission of alphanumerics by television. Displays, 2, 295-299.

WHITE, T.A., and REID, G.M. /‾1981_7, Quality of PAL colour television pictures impaired by random noise : stability of subjective assessment, IEE Proc, 128, Pt F, pp 231-236.

/‾1978-82_7 : a. 11/17 (EBU); b. 11/257 (France); c. 11/258 (France); d. 11/259 (France); e. 11/287 (UK); f. 11/288 (UK); g. 11/285 (UK); h. 11/292 (USA); i. 11/309 (Italy); j. 11/312 (France); k. 11/313 (France); l. 11/322 (Italy); m. 11/331 (German Democratic Republic); n. 11/343 (Japan), o. 11/357 (Italy); p. 11/71 (France).

/‾1974-78_7 : 11/360 (France).

ANNEX VI

SUBJECTIVE QUALITY OF ALPHANUMERIC AND GRAPHIC PICTURES

Different types of new services may be transmitted in television channels; they use digital codes to describe alphanumeric and graphic pictures (see Report 802). Some transmission parameters have an effect on the quality of displayed pictures : page resolution (number of rows per page and number of characters per row) in the case of alphamosaic coding of teletext, character cell resolution (number of pixels and lines per cell) in the case of DRCS coding, picture resolution in the case of broadcast audiography, facsimile or teletex. Further, the effects of transmission errors which may effect the codes should also be considered. Thus, measurements of quality and determinations of objective-to-subjective relationships for these parameters are necessary.

The contribution /¯CCIR, 1978-82a/ gives a detailed summary of the different aspects of quality assessment for these pictures which have characteristics distinct from those of conventional television pictures. Parameters such as pixel format, character cell resolution, spacings, colours, layout, ... have effects on various quality attributes : legibility, quality, comfort, annoyance, effort of reading, fatigue and aesthetic considerations. Two main aspects were discussed: the viewing conditions and the assessment methods.

1. Viewing conditions

Recommendation 500-1 defines viewing conditions for television pictures corresponding to low illumination levels in the room. It is likely that alphanumeric and graphic pictures would be viewed also in normal lighting conditions. Thus, a complementary set of viewing conditions is suggested for study : illumination of 500 lux, screen maximum luminance from 70 to 200 cd/m^2, screen contrast ratio from 30 to 50 and a value of ¼ for the ratio of background luminance (from the walls of the room) to maximum screen luminance. Viewing distance is also discussed (from 4 to 8 times picture height).

2. Assessment methods

A considerable number of studies have been made on typographical aspects. Most of them have used what might be called "performance measures" like detection or recognition thresholds, recognition ratio, speed of reading, etc. Very few have used "subjective measures" which are of a conventional use in television. It is felt that new systems transmitted in television channels should have good performance (for example, percentage of good recognition of letters higher than 95 %). Quality scale and impairment scale from Recommendation 500-1 could thus be used efficiently although studies are needed to establish the way in which these scales can be related to legibility. A comparison with speech quality assessment methods (CCITT) was tried and a 5-grade scale of "effort of reading" is suggested for studies.

The contribution /‾CCIR, 1978-82b_/ compares results of subjective assessments made using two different 5-grade scales given below :

Quality of legibility scale	Reading effort scale
- excellent legibility	- no reading effort
- good legibility	- attention necessary, but no appreciable reading effort
- fair legibility	- moderate reading effort
- poor legibility	- substantial reading effort
- bad legibility	- very substantial reading effort
(Quality of legibility scale)	(Reading effort scale)

The scale on the left concerns legibility, the one on the right, the reading effort (it was found important to make each wording of grades explicit). The mean values of the grades obtained with the reading effort scale are generally higher than those obtained with the legibility scale and the range actually used by the observers is higher in the case of the reading effort scale.

Another experiment /‾White, 1980; White, 1981_/ used the quality scale in Note 1 of Annex I of Report 405-3 (in accordance with Recommendation 500-1) to assess opinion of both overall quality and overall legibility of typescript transmitted by a television system of variable line standard and bandwidth. For each condition, two models, one of greater complexity and accuracy, but both invoking the concept of "Imp-scale" addition (Annex II), were discovered to describe the combined effects of limited horizontal and vertical definition. Legibility was also measured in terms of proportion of correctly identified characters. However, legibility in such terms remained high when quality was low, and it is evident that, usually, the former criterion is less useful.

Another study /‾Sallio, Morin, 1981_/ did comparisons of performance and subjective methods on text material using fixed-width characters and variable-width characters. Subjective methods were shown to be the more sensitive."

At the moment, too few experiments have been made on the subjective quality of alphanumeric and graphic pictures in the television field; it would be very useful if Administrations could make contributions on the subject.

<u>REFERENCES</u>

SALLIO, P. and MORIN, C. /‾Nov. 1981_/, Lisibilités et nouveaux services. Applications à l'analyse et à l'évaluation d'un paramètre graphique; matrices fixes, matrices variables. Colloque international sur la communication graphique et écrite, Paris.

WHITE, T.A. /‾1980_/ Transmission of alphanumerics by television : assessment of typescript by "experts". Proc. 9th International Symposium on Human Factors in Telecommunications, New Jersey, 24-34.

WHITE, T.A. /‾1981_/ Transmission of alphanumerics by television. Displays <u>2</u>, 295-299.

<u>CCIR Document</u> :

/‾1978-82_/ : a. 11/68 (France); b. 11/259 (France).

ANNEX VII

SPECIFICATIONS AND ALIGNMENT PROCEDURES FOR PICTURE SIGNAL SOURCES AND DISPLAYS

Differences in the alignment procedures and performance of the picture signal sources and the picture displays used in subjective tests may contribute to differences between the results obtained in such tests in different laboratories. Study Programme 3A-1/11 (MOD I), recognizing this problem, invites studies on the minimum performance specifications and on the alignment procedures that can be recommended for picture signal sources and for picture displays to be used in subjective tests.

The EBU has recommended to its member organizations /‾CCIR, 1978-82_7 a detailed alignment procedure for grade 1 colour monitors used in television production centres. Although referring to monitors used for a different purpose, this contribution is nevertheless relevant to the monitors used in subjective tests because both these applications require high-quality and reproducible performance standards. The document is therefore noted as a valuable preliminary input to the studies on picture sources and displays.

REFERENCES

CCIR Document

/‾1978-82_7 : 11/253 (EBU)

BIBLIOGRAPHY

EBU /‾1980_7 : "Procedure for the operational alignment of grade 1 colour monitors" EBU Technical Recommendation R25-1980.

ANNEX V

SUBJECTIVE PICTURE QUALITY IN DIGITAL TELEVISION SYSTEMS

Picture impairments in digital television systems are, in general, different from those that occur in analogue systems, and depend on the methods of coding and redundancy reduction employed. Consequently, in early laboratory studies of coding systems which seek to compare different systems or to optimize a particular system, it may be advantageous to use different testing techniques specially chosen to visually distinguish specific impairments more clearly.

Dynamic test pictures have advantages when testing digital and mixed analogue-digital-analogue systems. These test pictures should include linear and rotational movement as well as slow signal-amplitude fluctuations [CCIR, 1974-78f]; the latter fluctuations can reveal quantizing contour effects in almost uniform areas.

In some cases, subjective assessments may usefully be supported by instrumental methods. The "comparison-field" (e.g., split-screen) technique has been proposed for testing digital systems [Krivosheev, 1976]. In this method, comparison fields are displayed side by side with those areas of the original picture which are being studied. Two types of such fields are used: synthesized fields and natural fields. The first type is generated electronically, and by suitable choices of brightness, chromaticity, structure and configuration can correspond closely with the picture being assessed. Measurements of quantizing noise may be made using this method. The second type of comparison field consists of parts of the original picture; this type can, for example, be used when comparing the input to, and output from, a system.

APPENDIX B

LIST OF DOCUMENTS REVIEWED IN TASK ONE OF THIS CONTRACT

Documents Reveiwed in Part of Task One of this Contract

In carrying out Task One of this contract a considerable body of literature was identified, obtained and reviewed. It was found convenient to separate the documents obtained into the following headings:

1. - Documents Relating to the Establishment of Impairment Units for Video Systems

2. - Documents Relating to the Summation of Impairments of Video Systems

3. - Documents Relating to the Subjective Evaluation of Video Systems

4. - CCIR, CCITT and Other Regulatory Documents Applicable to the Use of Impairment Units (and Their Summation) in the Evaluation of Video Systems

5. - Other Documents Relevant to the Feasibility of the Use of Impairment Units, and Their Summation, in Evaluating NTSC Video Transmission Systems.

In this Appendix these documents are listed under the five headings given above in the chronological order of their publication.

All documents identified in this Appendix are held in the files of P.A. LAPP LIMITED. Those documents marked "*" are considered documents key to the findings of this study.

1. Documents Relating to the Establishment of Impairment Units for Video Systems

1.1 A.D. Fowler: Observer Reaction to Low Frequency Interference in Television Pictures. Proc. IRE, Vol. 39, Oct., 1951.

1.2 *A.D. Fowler: Observer Reaction to Video Crosstalk. Journal of the SMPTE, Vol. 57, Nov., 1951.

1.3 J.M. Barstow and H.N. Christopher: The Measurement of Random Monochrome Video Interference. Trans. Am Inst. Elect. Engineers, Part I, January, 1954.

1.4 J.M. Barstow and H.N. Christopher: The Measurement of Random Video Interference to Monochrome and Color Television Pictures. Trans. Am Inst. Elect. Engineers, Part I, 81, November, 1962.

1.5 W.K.E. Geddes: The Relative Impairment Produced by Random Noise in 405-Line and 625-Line Television Pictures, EBU. Review, Part A - Technical, No. 78, April, 1963.

1.6 *R.D. Prosser, J.W. Allnatt and N.W. Lewis: Quality Grading of Impaired Television Pictures. Proc. IEE, Vol. 111, No. 3, March, 1964.

1.7 B.W. Osborne, A.M. Peverett and D.A.R. Wallace: The K Rating of Television Networks. The Television Society Journal, Vol. 10, No. 10, 1964.

1.8 R.D.A. Maurice: The Specification of an Adequate Television Signal. The Television Society Journal, Vol. 11, No. 4. 1965.

1.9 J.W. Allnatt and R.D. Prosser: Subjective Quality of Television Pictures Impaired by Long-Delayed Echoes. Proc. IEE, Vol. 112, No. 3, March, 1965.

1.10 R.D. Prosser and J.W. Allnatt: Subjective Quality of Television Pictures Impaired by Random Noise. Proc. IEE, Vol. 112, No. 6, June 1965.

1.11 J.W. Allnatt: Subjective Quality of Colour-Television Pictures Impaired by Gain and Delay Inequalities Between the Luminance and Chrominance Channels. Proc. IEE, Vol. 112, No. 10, October 1965.

1.12 J.W. Allnatt and R.D. Prosser: Subjective Quality of Colour-Television Pictures Impaired by Random Noise. Proc. IEE, Vol. 113, No. 4, April 1966.

1.13 Yutaka Itoh: System Designing of Color Television Facilities. Journal of SMPTE, Vol. 76, No. 12, December 1967.

1.14 *J.W. Allnatt and E.J.W. Bragg: Subjective Quality of Television Pictures Impaired by Sinewave Noise and Low-Frequency Random Noise. Proc. IEE, Vol. 115, No. 3. March, 1968.

1.15 *L.E. Weaver: The Quality Rating of Colour Television Pictures. Journal of the SMPTE, Vol 77, June, 1968.

1.16 A. Brown: Television Co-Channel Interference with PAL Colour Transmissions. IEE Conference Publication, No. 46, Part 1, 1968.

1.17 J.W. Allnatt: Random Noise in the Reception of Colour Television. Electronic Engineering. Nov. 1968.

1.18 J.M. Corbett, J.R. Taylor and J.W. Allnatt: Subjective Quality of Colour-Television Pictures Impaired by Video Crosstalk. Proc. IEE, Vol. 116, No. 2, February, 1969.

1.19 John R. Cavenaugh: A Single Weighting Characteristic for Random Noise in Monochrome and NTSC Colour Television. Journal of the SMPTE, Vol. 79, Feb. 1970.

1.20 B.W. Osborne: Colour Television Signal Demodulation and the K Rating. The Relay Engineer, Vol. 9, No. 2, July, 1971.

1.21 *C.A. Siocos: Numerical Values for Subjective Picture Quality in Television. Journal of the SMPTE, Vol. 80, Dec. 71.

1.22 John W. Allnatt: Subjective Assesssment of Impairments in Television Pictures. Appl. Ergon, Vol. 2.4, 1971.

1.23 *D.E. Pearson: Methods of Scaling Television Picture Quality: A Survey In Picture Bandwidth Compression, T.S. Huang and O.J. Tretiak, Eds., New York: Gordon & Breach, 1972.

1.24 John Allnatt: Opinion Distribution Model for Subjective Rating Studies. Int. J. Man-Machine Studies (1973) 5.

1.25 J.M. Corbett and J.W. Allnatt: Subjective Quality of Television Pictures Impaired by Short-Delay Echo. Proc. IEE, Vol. 121, No. 7, July, 1974.

1.26 *W. Flor: Some Results of Evaluations by Control Stationary and Mobile T4-2. Provisional Report: ORF. Vienna, 8.10.1976.

1.27 Gillian M. Reid: Subjective Data for Random Noise Measured Through the New CMTT Unified Weighting Network for Television. Elec. Lett., Vol. 13, No. 18, Sept. 1977.

1.28 *P. Groenveld and C.A. de Beer: Subjective Evaluation of Still Crossmodulation Patterns in CATV Systems.
Cable TV Eng, Vol. 11, No. 3, Aug., 1978.

1.29 *Impairment Units and Objective Measurements.
IEC/SC 12G-WG3 (Street) 1, Aug. 1978.

1.30 Impairment Units and Objective Measurements
IEC/SC 12G-WG3 (Street) 1, Appendix. Sept. 1978.

1.31 *I.F. MacDiarmid and J.W. Allnatt: Performance Requirements for the Transmission of the PAL Coded Signal. Proc. IEE, Vol. 125, 1978.

1.32 Summation of Impairments - "Comments on IEC DOC 12A (Secretariat) 158": IEC/12G WG3 (Vahamaa). 1979.

1.33 *Robert A. Bednarek: On Evaluating Impaired Television Pictures by Subjective Measurements. IEEE Trans. on Broadcasting, Vol. BC-25, No. 2, June 1979.

1.34 *C.A. Siocos and G. Chouinard: Subjective Impairment Units in Relation with Oscilloscope Graticules for Evaluating Short-Time Linear Waveform Distortion of the Luminance Signal in 525-Line Television
IEEE Trans. on Broadcasting, Vol. BC-25. No. 2, June 1979.

1.35 *Impairment Units IEC Discussion Document.
IEC/SC 12G-WG3 (Secretary) 16, Nov. 1979.

1.36 Impairment Units - Letter: JN Corbett/GH Barlow:
IEC/SC 12G-WG3 (Secretary) 21.
March, 1980.

1.37 J.W. Allnatt: Subjective Assessment Method for Television Digital Codecs. Elec. Lett., Vol. 16. No. 12, June 5, 1980.

1.38 T.A. White and J.W. Allnatt: Double-Stimulus Quality Rating Method for Television Digital Codecs. Elec. Lett., Vol. 16, No. 18. Aug. 28, 1980.

2. Documents Relating to the Summation of Impairments of Video Systems

2.1 *N.W. Lewis and J.W. Allnatt: Subjective Quality of Television Pictures with Multiple Impairments. Elec. Letters, Vol. 1, No. 7, September 1965.

2.2 R.D.A. Maurice: Tolerances for PAL Colour Television. The Royal Television Society Jl., Vol. 12, No. 4, 1968.

2.3 N.W. Lewis and J.W. Allnatt: Subjective and Objective Impairments in Television Pictures. IEE Conference Publication No. 46, (1968).

2.4 C.A. Siocos: Subjective Impairment of Television Pictures - Considerations Regarding the Summability of Impairment Units. Journal of the SMPTE, Vol. 81, December 1972.

2.5 D.E. Pearson: Methods for Scaling Television Picture Quality: A Survey. In Picture Bandwidth Compression, T.S. Huang and O.J. Tretiak, Eds, New York: Gordon & Breach, 1972.

2.6 John Allnatt: Subjective Rating and Apparent Magnitude. Int. J. Man-Machine Studies (1975) 7.

2.7 *J.W. Allnatt: Opinion-Rating Model Applied to Television-Transmission-System Studies. Proc. IEE, Vol. 122, No. 8, Aug. 1975.

2.8 W. Flor: Quality Grading Models for Multiple Impaired Television Pictures. Technical Paper T4.23 (ORF) July 1977. (Original in German)

2.9 J.W. Allnatt: Subjective Quality of Television Pictures Impaired by Linear Waveform Distortion. Elec. Letters, Vol. 13, No. 18. September 1977.

2.10 Robert A. Bednarek: On Evaluating Impaired Television Pictures by Subjective Measurements. IEEE Trans. on Broadcasting, Vol. BC-25, No. 2. June, 1979.

2.11 J.S. Goodman and D.E. Pearson: Multidimensional Scaling of Multiply-Impaired Television Pictures. IEEE Trans. on Syst., Man., & Cyber., Vol. SMC-9, No. 6. June, 1979.

2.12 *John Allnatt: Television Measurements Through Psychophysics to Subjective Picture Quality. Rad & Elec. Eng., Vol. 49, No. 12, Dec., 1979.

3. Documents Relating to the Subjective Evaluation of Video Systems (All in volume 2 except as noted.)

3.1 Relationships Between Subjective and Objective Parameters - Extract from BPO RDR 691. (not dated)

3.2 P. Mertz, A.D. Fowler and H.N. Christopher: Quality Rating of Television Images. Proc, IRE, Vol. 38, Nov., 1950.

3.3 L.E. Weaver: Subjective Impairment of Television Pictures. Elec. and Radio Eng. May, 1959.

3.4 Charles E. Dean: Measurements of the Subjective Effects of Interference in Television Reception. Proc. IRE, Vol. 48, June 1960.

3.5 Hans Schmid: Measurement of Television Picture Impairments Caused by Linear Distortions. Journal of SMPTE, Vol. 77, March 1968.

3.6 *J.M. Corbett: Effect of Observer Adaptation on the Results of Television Quality-Grading Tests. Proc. IEE, Vol. 117, No. 3, March 1970.

3.7 J. Cavanaugh and A.M. Lissman: Subjective Effects of Differential Gain and Differential Phase Distortions in NTSC Colour Television Pictures. Journal of the SMPTE. Vol. 80, Aug. 1971.

3.8 *J.W. Allnatt and J.M. Corbett: Adaptation in Observers During Television Quality-Grading Tests. I: Adaptation as a Function of the Conditioning Situation. Ergonomics, Vol. 15, No. 4, 1972.

3.9 *J.W. Allnatt and J.M. Corbett: Adaptation in Observers During Television Quality-Grading Tests. II: Progress of Adaptation During an Experiment. Ergonomics, Vol. 15, No. 5, 1972.

3.10 A.M. Lessman: The Subjective Effects of Echoes in 525-Line Monochrome and NTSC Colour Television and the Resulting Echo Time Weighting. Journal of the SMPTE, Vol. 81, December 1972.

3.11 *J.R. Cavanaugh and A.M. Lessman: The Subjective Effect of Random Noise Spectra on 525-Line NTSC Colour Television. Journal of the SMPTE, Vol. 83, No.10, Oct., 1974.

3.12 A.S. Taylor: TV Picture Interference Study. National Science Foundation. May, 1977.

3.13 P.L. MAMBO: Measurements and Subjective Evaluation of Picture Quality in CATV Systems. TP Secs 77.3, Carleton University, Aug. 1977.

3.14 T.A. White: Subjective Assessment of Interline Flicker in Television Displays - A First Approach. Proc. Eighth Int. Symp. on Human Factors in Telecomms, Sept., 1977.

3.15 P.L. Mambo & D.C. Coll: Perceived Picture Quality in CATV Systems with Impairments. IEEE Trans. on Cable Telev., Vol. CATV-4, No. 1, Jan., 1979.

3.16 Shigeru Yamazaki and Yukio Endo: Quantitative Measurement of a Subjective Evaluation of TV Pictures Impaired by Multiple "Ghosts". NHK Lab Note, No. 261, Mar., 1981.

3.17 Paul Kaiser and Gareth McIntosh: Subjective Assessment of Picture Quality. CCTA Official Technical Record. 1982.

3.18 A. Lymer and W. Gosling: Determination of A.M. Co-channel Protection Ratio For Land Mobile Signals by Means of Subjective Testing. Rad. and Elec. Eng., Vol. 52, No. 1. January, 1982.

4. <u>CCIR, CCITT and Other Regulatory Documents Applicable to the Use of Impairment Units (and Their Summation) in the Evaluation of Video Systems</u>

4.1 CCIR Rec. 421-1: Requirements for the Transmission of Television Signals Over Long Distances (System I Excepted). XIth Plenary Assembly, Oslo, 1966, Vol. V.

4.2 CCIR Rep. 405: Subjective Assessment of the Quality of Television Pictures. XIth Plenary Assembly, Oslo, 1966, Vol. V.

4.3 CCIR Rep. 405-1: Subjective Assessment of the Quality of Television Pictures. XIIth Plenary Assembly, New Delhi, 1970, Vol. V, Part 2.

4.4 CCIR Rec. 418-2: Ratio of the Wanted-to-Unwanted Signal in Monochrome Television. XIIIth Plenary Assembly, Geneva, 1974, Vol. XI.

4.5 CCIR Rep. 405-2: Subjective Assessment of the Quality of Television Pictures. XIIIth Plenary Assembly, Geneva, 1974, Vol. XI.

4.6 CCIR Rec. 451-2: Requirements for the Transmission of Television Signals Over Long Distances (System 1 Only). XIIIth Plenary Assembly, Geneva, 1974, Vol. XII.

4.7 CCIR Rep. 486-1: Transmission Performance of Television Circuits Designed for Use in International Connections. XIIIth Plenary Assembly, Geneva, 1974.

4.8 CCIR Rec. 500: Method for the Subjective Assessment of the Quality of TV Pictures. XIIIth Plenary Assembly, Geneva, 1974.

4.9 CCIR Draft Rec. 500(Rev.76): Method for the Subjective Assessment of the Quality of Television Pictures. XIVth Plenary Assembly, Kyoto, 1978.

4.10 CCIR Rep. 405-3: Subjective Assessment of the Quality of Television Pictures. XIVth Plenary Assembly, Kyoto, 1978, Vol. XI.

4.11 CCIR Rep. 624-1: Characteristics of Television Systems. XIVth Plenary Assembly, Kyoto, 1978, Vol. XI.

4.12 CCIR DRAFT Rec. 567 (MOD F): Transmission Performance of Television Circuits Designed for Use in International Connections. XVth Plenary Assembly, Geneva, 1982.

4.13 *CCIR DRAFT Rep. 405-3 (MOD F): Subjective Assessment of the Quality of Television Pictures. XVth Plenary Assembly, Geneva, 1982.

4.14 *CCIR DRAFT Rep. 313-4 (MOD F): Assessment of the Quality of Television Pictures. XVth Plenary Assembly, Geneva, 1982.

4.15 *CCIR DRAFT Rec. 500-1 (MOD F): Method for the Subjective Assessment of the Quality of Television Pictures. XVth Plenary Assembly, Geneva, 1982.

4.16 CCIR DRAFT Rep. AC/11 (MOD F): Experimental Results Relating Picture Quality to Objective Magnitude of Impairment. XVth Plenary Assembly, Geneva, 1982.

4.17 *CCIR DRAFT Rep. AR/11: Subjective Quality of Television Pictures in Relation to the Main Impairments of the Analogue Composite Television Signal. XVth Plenary Assembly, Geneva, 1982.

5. <u>Other Documents Reviewed as Being of Possible Interest to the Feasibility of the Use of Impairment Units, and Their Summation, in Evaluating NTSC Video Transmission Systems</u>

5.1 Consumer Electronics Group: Consumer Electronics Service Technician: Interference Handbook. CES Publication. (not dated)

5.2 Pierre Mertz: Influence of Echoes on Television Transmission. Journal of SMPTE, Vol. 60, May 1953.

5.3 Johannes Muller: Uber den Zusammenhang von Einschwingverhalten und Bildgute bei Fernsehubertragungssystemen. FTZ, July, 1953.

5.4 N.W. Lewis: Waveform Responses of Television Links Proc. IEE, Vol. 101, 1954.

5.5 W.N. Sproson, S.N. Watson and M. Campbell: The BBC Colour Television Tests: An Appraisal of Results. BBC Engineering Monograph No. 18, May, 1958.

5.6 B.W. Osborne: Picture Quality Assessment and Waveform Distortion Correction on Wired Television Systems. J. Brit. IRE, Vol. 23, May, 1962.

5.7 D.N. Carson: CATV Amplifiers: Figure of Merit and the Coefficient System. IEEE Trans on Commun. Tech., Vol. COM-14, No. 4, August, 1966.

5.8 Peter Wolf: Modification of the Pulse-and-Bar Test Signal with Special Reference to Application in Colour Television. Journal of SMPTE, Vol. 75, January, 1966.

5.9 C.A. Siocos: Vertical Interval Test and Reference Signals (VITS) in the CBC Television Network. Journal of SMPTE, Vol. 75, No. 2. February, 1966.

5.10 C.A. Siocos: Chrominance-to-Luminance Ratio and Timing Measurements in Color Television. IEEE Trans. on Broadcasting, Vol. BC-14, No. 1, Mar., 1968.

5.11 Thomas M. Gluyas: Correlation Between TV Transmitter Performance Measurements and Picture Quality. IEEE Trans on Broadcasting, Vol. BC-14, No. 1, Mar., 1968.

5.12 Robert E. Mallon: Application of K-Rating to U.S.A. NTSC Systems. Journal of SMPTE, Vol. 79. January, 1970.

5.13 G.C. Higgins: Methods for Analyzing the Photographic System, Including the Effects of Nonlinearity and Spatial Frequency Response.
Photographic Science and Engineering, Vol. 15, No. 2, March-April, 1971.

5.14 Daan M. Zwick: Television Receiver White Color: A Comparison of Picture Quality with White References of 9300 K and D6500. Journal of SMPTE, Vol. 82. April, 1973.

5.15 B.S. Rao: Carrier-to-Noise Ratio - The Prime Determinant of Picture Quality in FM Systems. AEU, No. 2, 1973.

5.16 Stanley M. Soliday and James A. Gardner: Picture Quality Judgements in a Digital Television System. Human Factors, 1974, 16(2).

5.17 *Lucien M. Biberman: Fallacy and Fact of Sampled Imagery Displays. Human Factors, 1974, 16(3).

5.18 B. Arnold: Required System Triple-Beat Performance. R.C.A. Engineer, Vol. 19, Dec.73/Jan.74.

5.19 J.W. Allnatt and J.M. Corbett: Comparisons of Category Scales Employed for Opinion Rating. Proc. IEE, Vol. 121, No. 8, August, 1974.

5.20 T.M. Straus: The Relationship Between the NCTA, EIA, and CCIR Definitions of Signal-to-Noise Ratio. IEEE Trans. on Broadcasting, Vol. BC-20, No. 3, Sept., 1974.

5.21 Masayuki Inoue and Yukio Kobayashi: Single Frequency Disturbance in Video Telephones. Rev. Electr. Commun. Lab. (Japan), Vol. 23, No. 1-2, Jan-Feb, 1975.

5.22 Determination of Echo-Signal Tolerances in Transmission Channels of Full SECAM Signals. Tekhnika kino i televideniya, (9) pp 62-68. 1976.

5.23 B.W. Osborne: The Assessment of Picture Quality on Cable Television Systems by Means of Engineering Audits. Cable Television Engineering, Vol. 10, No. 8. April, 1976.

5.24 K. Lucas: The Numerical Basis for ORACLE Transmission. IBA Tech. Rev., No. 9, 1976.

5.25 Robert W. Edmonds: The Use of Echo Time-Weighting to Derive Oscilloscope Graticules for Rating Television Transmission Performance. J. SMPTE, Vol. 85, June, 1976.

5.26 Lorimer Clayton: FM Television Signal-to-Noise Ratio. IEEE Trans on Cable Telev., Vol. CATV-1, No. 1, Oct. 1976.

5.27 R.D.A. Maurice: Developments in the Pal Colour-Television System: A Unified System of Tolerances. Proc. IEE, Vol. 125, No. 6, June, 1978.

5.28 Shri K. Goyal and Stephen C. Armfield. Reception of Teletext under Multipath Conditions. IEEE Trans. on Cons. Electr., Vol. CE-25, No. 3. July, 1979.

5.29 N.S. Copley: World Standards Conversion. Video Systems, Oct., 1979.

5.30 John M. Hood: Perceptibility of Composite Triple Beat. CCTA Technical Records, 1980.

5.31 Shigeru Yamazaki and Yukio Endo: A Quantitative Measurement of Subjective Effects of TV Multiple Ghost Images by 'Perceived D U Ratio'. IEEE Trans. on Broadcasting, Vol. BC-26, No. 3, Sept., 1980.

5.32 Theodor Bossert: Institut für Rundfunktechnik GMBH. NTG-Fachberichte Vol. 72, 30.10.1981.

5.33 Technical Standards and Procedures for Broadcasting Receiving Undertakings (Cable Television). Broadcast Procedure 23, Issue 2, January 1, 1982.

5.34 André Pomerleau, Denis Tardif et Roger Brousseau: Mesure quantitative de la qualité d'une image de télévision. Université Laval, Rapport No. LT-81-8297. Feb. 1982.
