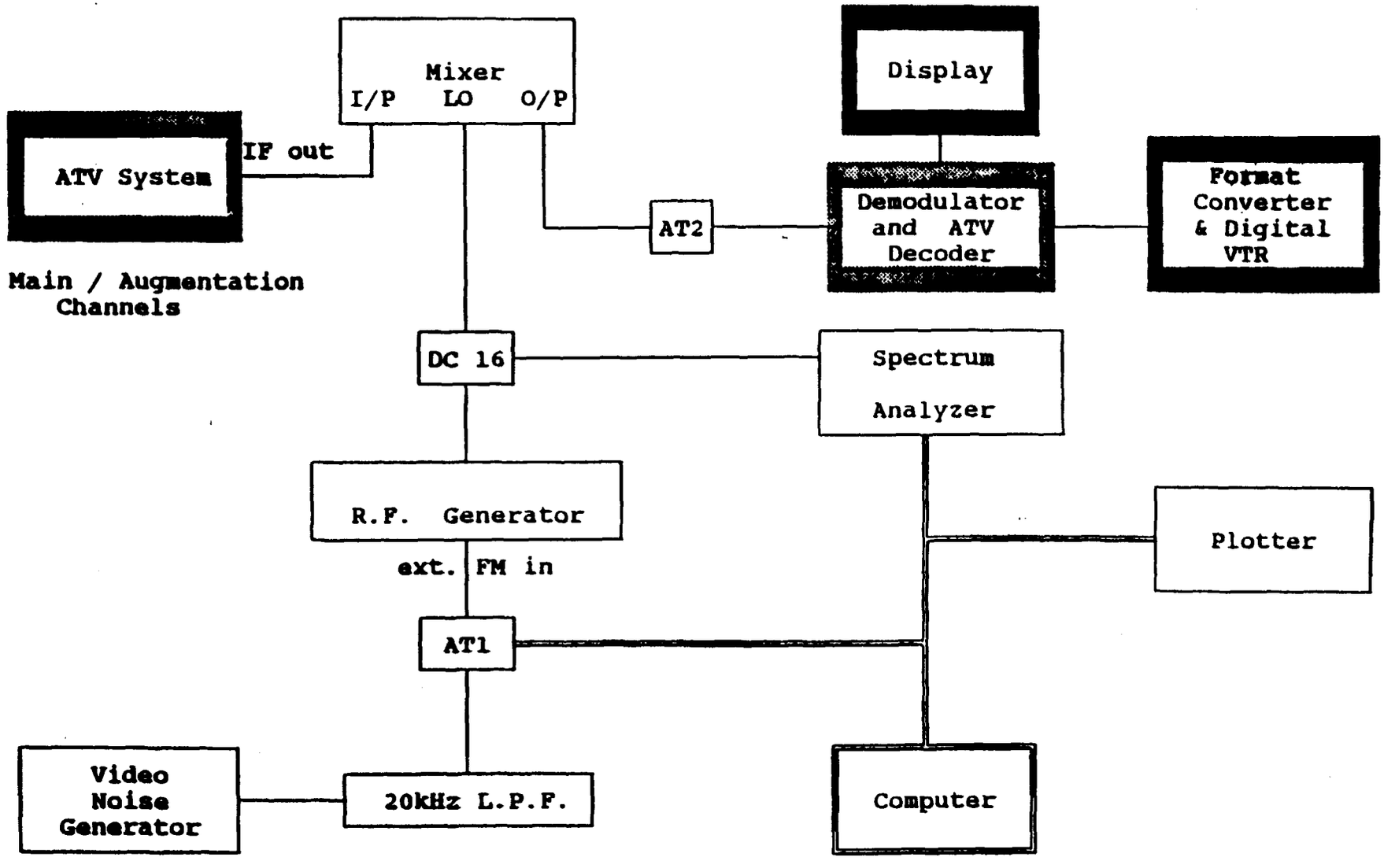


Figure 7.1 Incidental Carrier Phase Modulation Equipment Setup Block Diagram.

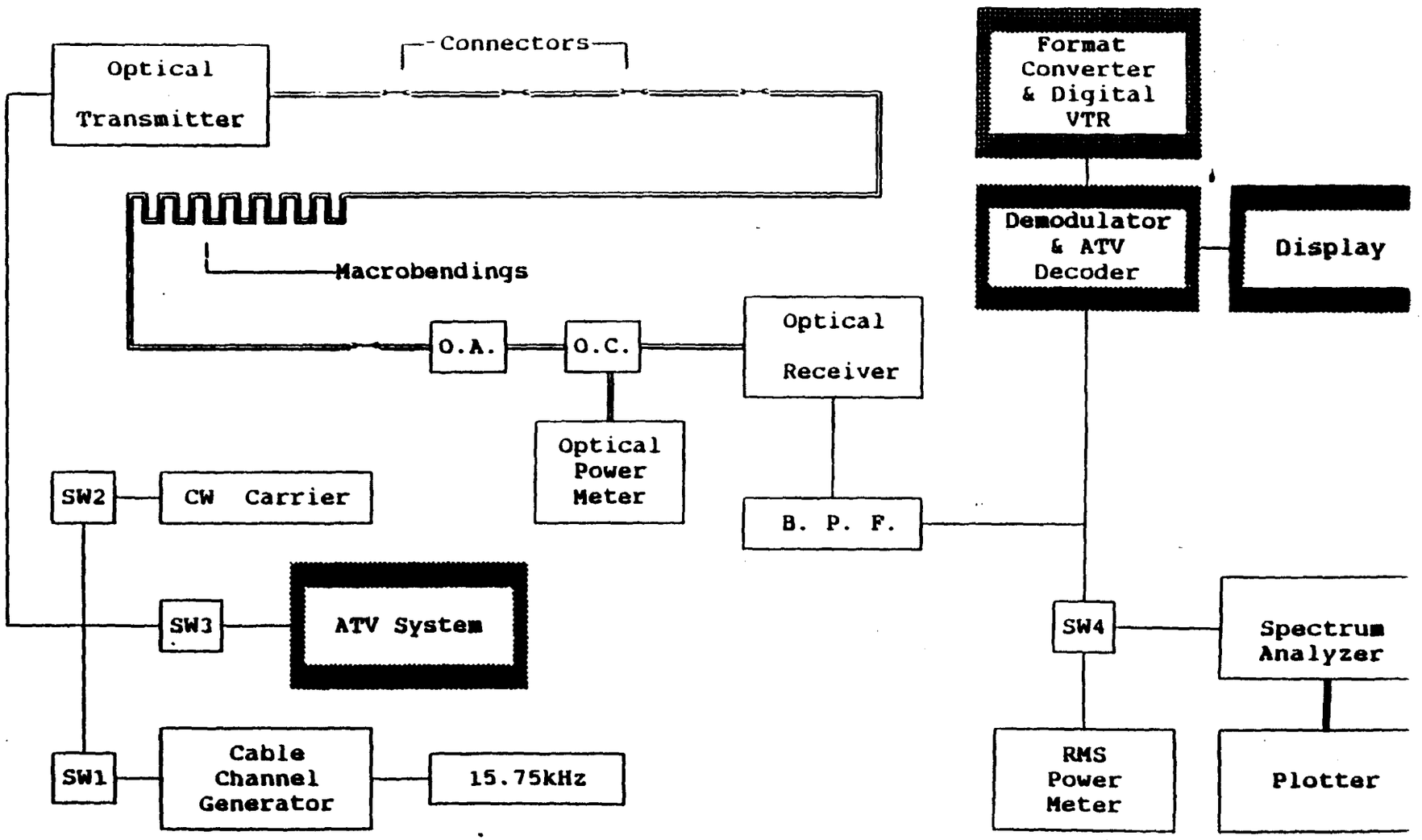


Figure 8.1 Optical Transmission Test Equipment Setup Block Diagram.

SSWP2-0390
6 Sep 90
Rev 28 Sep 90

ATV Test Procedures
Video Subjective Tests

TABLE OF CONTENTS

Page

8. TEST METHODS FOR INDIVIDUAL IMPAIRMENTS

1. INTRODUCTION

1.1 Relationship of Video Subjective Test Plan to Video Objective Test Plan

This test plan is meant to be used in conjunction with other FCC Advisory Committee documents.

1.2 General Description of Test Types

There are three general types of subjective tests, ranging (which includes finding thresholds), rating, and expert observation and comment. The purpose of ranging is to establish the threshold of visibility of an impairment, the point where the impairment renders the signal unusable, and some of the steps in between. During rating tests a group of non-expert observers watch television pictures and decide on their quality or on the effect of an impairment. During an expert observation and comment test, expert observers watch television pictures and offer non-quantitative comments on the pictures.

1.3 Ranging Tests-General Description

Ranging tests fall into several categories which are described below.

All ranging tests do have several factors in common:

A. Expert observers look at a single large screen video display or at a bank of NTSC receivers.

B. A single ranging test may be part of a larger ensemble of ranging tests for a given variable, e.g., co-channel interference may be measured at several receiver input power levels.

C. The use of audible cues to signal changes in impairment level will be investigated during the prove-in tests with NTSC.

D. Immediately after finding the endpoints (threshold of visibility and point of unusability), the observers will determine the increments to be used in possible future rating tests.

1.3.1 Single Variable ATV Ranging Test

In single variable ATV ranging the value of the single impairment being tested is slowly varied. The expert observers then decide, as described elsewhere [See Section 1.6] on the threshold of visibility and the point of unusability.

The threshold of visibility and the point of unusability make up a range of impairment which then is divided into steps. Certain of those steps will be selected by the experts for use in possible subsequent rating tests.

1.3.2. Multiple Variable ATV Ranging Test

In multiple variable ATV ranging it is recognized that many impairments are characterized by more than one variable, e.g., multipath (Undesired signal level and delay), airplane flutter (Undesired level, delay, and Doppler shift). For these impairments it is necessary to hold all variables but one constant for any given ranging test. This means that there will be a large number of ranging tests for multiple variable ranging tests.

1.3.3. NTSC Calibration

In these tests the expert observers will ascertain when the impairment being investigated is visible on half (12) of the NTSC receivers. The level of the Undesired signal at which half the NTSC receivers show the impairment will constitute the median value of the threshold of visibility for Undesired > NTSC.

The 3/8/24 method, described below, is used.

An NTSC calibration is always done prior to the testing of ATV-into-NTSC interference or the testing of impairments or interferences to an Enhanced-NTSC signal as viewed on an NTSC receiver.

1.3.4. NTSC Ranging (3/8/24 Method)

In this method three observers watch 24 NTSC receivers arranged in three banks of eight receivers. In front of each bank of eight receivers there is a chair for an observer.

The impairment is increased until it is plainly visible on all receivers or until the maximum power level of the test apparatus has been reached. At this point each observer takes a chair in front of a receiver bank. Once the observers are seated, the impairment level is decremented. The observers then record upon which receivers the impairment is still visible. When all observers have checked their eight receivers and recorded the results, the impairment is decremented, and the process repeated. The process is repeated until none of the receivers shows the impairment.

After the above, the observers change chairs and repeat the above twice more, each time with an observer in front of a different set of receivers.

Having done the above procedure, the point of unusability is found by increasing the impairment until at least 2 of the three observers feel that the picture is unusable on half on the NTSC receivers.

1.4. Expert Observation and Comment-General Description

Expert observation and comment is a simple test wherein the expert observers watch a display(s) and provide written non-quantitative comments on what they see.

1.5. Rating Tests-General Description

The purpose of rating tests is to ascertain subjective assessments by non-expert observers of picture quality and impairments. These judgments will be made from video tapes played back for groups of non-expert observers.

There are two types of rating tests: quality and transmission impairment. Quality tests assess the intrinsic picture reproduction of a system under test relative to a reference television picture. Transmission impairment tests allow observers to rate the degree of degradation of pictures subjected to various levels of impairment (such ratings being made relative to the same picture unimpaired).

1.5.1. Quality Rating Test

The purpose of the quality tests is to ascertain impressions of picture quality. In these tests, non-expert observers will compare the unimpaired picture quality of the system under test with the quality of an unimpaired 1125/59.94 reference picture.

The actual method of testing is a variant of the double-stimulus continuous quality method described in CCIR Sec Rec 500-3. In this method the reference picture and the same picture via the system under test are shown sequentially to the observers. This is then repeated. Each observer then individually rates the reference and the system under test.

At least twenty non-expert observers shall view and rate the trials for each ATV system.

1.5.2. Transmission Impairment Rating Test

The purpose of the transmission impairment tests is to ascertain observer reaction to transmission impairment. In these tests the observers will compare an unimpaired picture with an impaired one (both pictures being in the format of the system under test).

The actual method of testing is the double-stimulus impairment method described in CCIR Rec 500-3. In this method the unimpaired picture at the "Desired" signal level being investigated is first shown to the observers followed by the same picture with the impairment added. Each observer then rates the system under test with respect to the unimpaired first picture.

At least twenty non-expert observers shall view and rate the trials for a given impairment and ATV system.

1.6. ATV Ranging Tests-Detailed Information

1.6.1. Interaction Among Observers

Each observer's vote on the threshold of visibility and on the point of unusability is recorded secretly so that the other observers do not know his decision. Discussion among observers is expected.

1.6.2. Determination of the ATV Threshold of Visibility

This is the method by which the threshold of visibility is ascertained. The method for one trial is as follows:

The impairment level is set at a level well below the threshold of visibility.

The level is then increased in 1 db steps every three seconds. (Depending on the impairment, the step size and timing may have to be modified.) As the impairment become visible to each observer, he presses his voting button. When four of the five observers have so voted, the level is recorded, and the impairment begins to decrement in 1 db steps. Two short beeps will be sounded to alert the observers that decrementing has begun. When four of the five observers vote that the impairment is no longer visible, the level is recorded, two short beeps are sounded and the impairment begins to increment in 1 db steps. This cycle is repeated four times for a total of five cycles.

If the maximum and minimum levels of the last cycle are within 3 db of each other, then the threshold of visibility is computed by averaging the last two maximum and minimum

levels.

If the maximum and minimum levels of the last cycle are not within 3 db then convergence must be verified as follows:

> The sum of the differences between the maximums and minimums for cycles two and three are compared to the sum of the differences of the maximums and minimums for cycles four and five. If the cycle four and five sum is smaller, then the threshold of visibility is the average of the maximum and minimums of cycles four and five. If the cycle two and three sum is smaller, then proper convergence has not been achieved, and the entire process must be repeated.

If the threshold of visibility persistently will not converge, an average of cycles four and five shall be taken and the fact of poor convergence noted.

1.6.3. Determination of the ATV Point of Unusability

The point of unusability found by taking the average of three iterations of the following:

> The impairment is set randomly at a point well below the point of unusability. It is then incremented until four of the five observers vote that the point of unusability has been reached.

On the second and third iterations, the impairment starting points are chosen to be random and different for each iteration.

1.6.4. Choice of Picture Material

For determination of the threshold of visibility, the video used for the Desired signal shall be chosen to be sensitive to the impairment being investigated. Generally a 50% gray flat field will be used. (The actual test pattern for a Desired ATV signal may have a dynamic characteristic. For the case of ATV-into-ATV interference testing, only a portion of the image can be a flat gray field.)

For determination of the point of unusability, a flat gray field may not be suitable. The preferred image contains features of interest (*e.g.*, colored text or graphics) which become indistinguishable or grossly distorted at the "point of unusability." (The test pattern for a Desired ATV signal may have a dynamic characteristic.)

The video used for the Undesired signal shall be chosen to maximally interfere with the Desired signal.

1.6.5. Determination of ATV Rating Test Steps

After the observers determine the threshold of visibility and the point of unusability, the test administrator shall inform the observers of the range between the two values. The observers shall then decide by consensus using the Method of Autoranging (described below) what test step sizes are to be used in subsequent rating tests. This information will be recorded by the test administrator as part of the test results.

The step sizes shall be determined by the Method of Autoranging. First a test picture [one of the still pictures provided by the FCC Advisory Committee (PS/WP-6) as impairment rating test material] replaces the test pattern as the desired signal. The

difference in dB between the threshold of visibility (TOV) and the point of unusability (POU) is divided into equal dB steps. The number of steps is selected by the observers based on their experience, the linearity of the impairment, and magnitude of the difference between TOV and POU.

The observers are then shown each of the Undesired levels resulting from their selection of the number of steps. The experts can either approve the quantity of steps selected, discard one or more Undesired levels because they feel a particular step does not reveal any significant change in impairment, or change the size of the steps.

From the above procedure, the observers shall select 6 points. These points shall include a step below threshold (impairment invisible), the step just below the point of unusability, and four intermediate steps which favor the low impairment end of the range which is the range of greatest interest.

1.7. NTSC Calibration-Detailed Information

The purpose of NTSC Calibration is to verify that the expert observers are getting test results that correspond to previous testing. This is done by conducting a test of the parameter being investigated using only NTSC.

The same method used for NTSC Ranging (3/8/24 method described below) is used for the NTSC Calibration.

1.8. NTSC Ranging Tests-Detailed Information

The following procedures are to be used for all testing of ATV-into-NTSC interference and for determining the threshold of visibility and point of unusability of Enhanced-NTSC signals as viewed on NTSC receivers.

1.8.1. Determination of the NTSC Threshold of Visibility - (3/8/24 Method)

In this method there are three observers and 24 NTSC receivers arranged in three banks of eight receivers. In front of each bank of eight receivers there is a chair for an observer.

The impairment is increased until it is plainly visible on all receivers or until the maximum power level of the test apparatus has been reached. At this point each observer takes a chair in front of a receiver bank. At each observer's chair there is a means for recording on which of the eight receivers in front of the chair the impairment is visible. Once the observers are seated, the impairment level is decremented. (Normally the decrements will be one dB, but depending on the impairment being investigated, the decrement size may be changed.) The observers then record upon which receiver(s) the impairment has become invisible at that level. When all observers have checked their eight receivers and recorded the results, the impairment is decremented, and the process repeated. The process is repeated until none of the receivers shows the impairment.

After the above, the observers change chairs and repeat the above twice more, each with an observer in front of a different set of receivers.

Using the results from the above procedure, the median NTSC receiver is selected (for the purposes of this procedure, the median receiver is the 12th receiver upon which the impairment is just visible as the level of impairment is decremented.) The Undesired

level at which the impairment is just visible on the median receiver is the "median threshold of visibility."

1.8.2. Determination of NTSC Point of Unusability

Having done the above procedure, the point of unusability is found by increasing the impairment until at least 2 of the three observers feel that the picture is unusable on half of the NTSC receivers. This level of impairment is the "median point of unusability."

1.8.3. Choice of Picture Material

For determination of the threshold of visibility, the video used for the Desired signal shall be chosen to be sensitive to the impairment being investigated. Generally a 50% gray flat field will be used.

For determination of the point of unusability, a flat gray field may not be suitable. The preferred image contains features of interest (*e.g.*, colored text or graphics) which become indistinguishable or grossly distorted at the "point of unusability."

The video used for the Undesired signal shall be chosen to maximally interfere with the Desired signal.

1.8.4. Determination of NTSC Rating Test Steps

After the observers determine the threshold of visibility and the point of unusability, the test administrator shall inform the observers of the range between the two values. The observers shall then decide by consensus using the Method of Autoranging (described below) what test step sizes are to be used in subsequent rating tests. This information will be recorded by the test administrator as part of the test results.

For this procedure, the experts must view a large-screen NTSC display, identical to the display to be used for the subsequent non-expert viewing.

The step sizes shall be determined by the Method of Autoranging. First a test picture [one of the still pictures provided by the FCC Advisory Committee (PS/WP-6) as impairment rating test material] replaces the test pattern as the desired signal. The difference in dB between the threshold of visibility (TOV) and the point of unusability (POU) is divided into equal dB steps. The number of steps is selected by the observers based on their experience, the linearity of the impairment, and magnitude of the difference between TOV and POU.

The observers are then shown each of the Undesired levels resulting from their selection of the number of steps. The experts can either approve the quantity of steps selected, discard one or more Undesired levels that they feel do not reveal any significant change in impairment, or change the size of the steps.

From the above procedure, the observers shall select 6 points. These points shall include a step below threshold (impairment invisible), the step just below the point of unusability, and four intermediate steps which favor the low impairment end of the range which is the range of greatest interest.

1.9. Quality Rating Tests-Detailed Information

There are four categories of quality rating tests:

ATV Basic Quality

Reference: 1125/59.94 or 1125/60
Test: ATV unimpaired

NTSC Reception Quality (Enhanced-NTSC systems only)

Reference: NTSC unimpaired
Test: NTSC received via unimpaired ATV

ATV Cable Quality

Reference: ATV unimpaired
Test: ATV via cable

ATV Fiber Quality

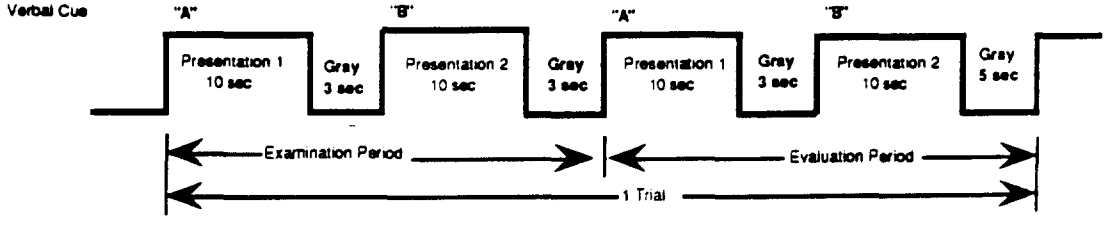
Reference: ATV unimpaired
Test: ATV via fiber

1.9.1. Test Material

The test material, selected by the FCC Advisory Committee, shall consist of still and moving pictures in a pseudo-random order. The same two pseudo-random orderings shall be used for all proponents. Each of the two pseudo-random sequences shall be such that the same picture material is never used in two consecutive trials. All test trials shall be presented twice.

1.9.2. Description of a Single Trial

A single trial consists of two presentations of a given piece of test material. One presentation is the reference, although it is not identified to the observers. The other presentation is the test, showing the same test material. The two presentations are shown twice separated by a period of mid-level gray. The following diagram shows the whole sequence:



Each presentation is identified to the observers by a recorded announcement saying either "A" or "B". On half the trials, the reference will follow the "A" announcement, and the system under test will follow the "B" announcement. On the other trials, the system under test will follow "A", and the reference will follow "B".

During a session the order of presentation of the reference and the system under test (e.g., which is "PRESENTATION 1" and which is "PRESENTATION 2") will change in a pseudo-random fashion. The same two pseudo-random orderings will be used for all proponents. The orderings shall be specified to the testing laboratory by the FCC Advisory Committee.

Each trial will be identified by a spoken cue, "Trial xx", recorded on the video tape.

1.9.3. Test Session Description

Prior to the beginning of a test session, formal instructions are given to the observers. These instructions are both written and oral. At this point the trials begin.

For a given ATV system a test block shall consist of four demonstration trials followed by five practice trials (The results from these five trials are not used.) followed by the actual trials.

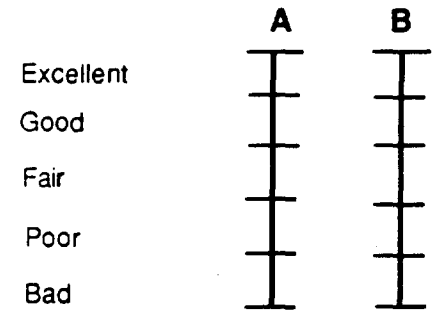
Each observer shall participate in only a single session.

The duration of the session, including a ten-minute break, is about sixty minutes.

Normally a session shall be conducted with up to five observers.

1.9.4. Rating Scale

The figure below shows the printed scale that the observers will use to rate each trial.



1.9.5. Instructions to Observers

The following representative instructions are given both in writing and announced via recording to each group of observers:

"This experiment is one of a series being carried out to evaluate new kinds of television for reception in the home. In this experiment, we ask you to evaluate the overall quality of the pictures you see."

"The experiment will consist of a series of evaluation trials. Each trial will involve...."

"We will now show you four sample trials. The first two of these will show still pictures, and the second two will show moving pictures. These will illustrate the kinds of pictures you will be asked to evaluate."

[PRESENT DEMONSTRATION TRIALS HERE]

"In the experiment, we will ask you to judge the overall quality of the pictures you see. To do this, you.... (Describe rating scale and judgment task)"

"Have you any questions?"

[AFTER QUESTIONS, BEGIN PRACTICE AND ACTUAL TRIALS]

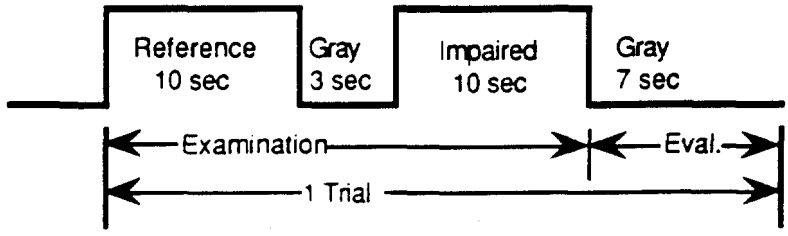
1.10. Transmission Impairment Rating Tests-Detailed Information

1.10.1. Test Material

The test material, selected by the FCC Advisory Committee, shall consist of still and moving pictures in pseudo-random order. For a given number of steps the same two pseudo-random orderings shall be used for all proponents. Each of the two pseudo-random sequences shall be such that the same picture material is never used in two consecutive trials. All test trials shall be presented twice. The first actual trial on the video tape shall have readily discernable impairment.

1.10.2. Description of a Single Trial

A single trial consists of two presentations of a given piece of test material. The first presentation shows the test material on the system under test without impairment at the Desired signal level under investigation. This is the reference. The second presentation is the system under test at the Desired signal level under investigation showing the same test material with impairment. The two presentations are shown once separated by a period of mid-level gray. The following diagram shows a complete trial:



The first member of the pair is identified to the observers as the unimpaired reference; the second member of the pair is identified as the test picture.

Each trial is identified by a spoken cue recorded on the video tape as "Trial xx".

For a given number of steps the same two pseudo-random orderings of the chosen levels of impairments and related pictures shall be used for all proponents for a given impairment. The orderings shall be specified to the testing laboratory by the FCC Advisory Committee. The first presentation for any test shall be the median impairment level identified in ranging tests.

1.10.3. Test Session Description

Prior to the beginning of a test session, formal instructions are given to the observers. These instructions are both written and oral. At this point the trials begin.

For a given impairment a test block shall consist of four demonstration trials followed by five practice trials followed by the actual trials each of which shall be repeated. [This allows the repeatability of the observers to be checked.] The number of actual trials is

vary depending on the number of steps selected during the ranging process.

Each observer shall participate in only a single session.

Normally a session shall be conducted with up to five observers.

On the average a test block for a given impairment will last about 30 to 35 minutes. To best utilize the non-expert observers it is desirable to have sessions which last from one to one and a half hours. Therefore two to three impairment blocks would normally be possible in a given session. A break lasting ten minutes is placed between blocks.

1.10.4. Rating Scale

The figure below shows the printed scale that the observers will use to rate each trial.

circle one

5 Imperceptible
4 Perceptible, but not Annoying
3 Slightly Annoying
2 Annoying
1 Very Annoying

1.10.5. Instructions to Observers

The following representative instructions are both given in writing and announced via recording to each group of observers:

"This experiment is one of a series being carried out to evaluate new kinds of television for reception in the home. In this experiment, we ask you to evaluate the effects of certain kinds of impairments that can occur in television reception."

"The experiment will consist of a series of evaluation trials. Each trial will involve

"We will now show you four sample trials. The first two of these will show still pictures, and the second two will show moving pictures. These will demonstrate the kinds of impairments you may see in the experiment to follow."

[PRESENT DEMONSTRATION TRIALS HERE]

"In the experiment, we will ask you to judge the degree of impairment to the pictures you will see. To do this, you.... (Describe rating scale and judgment task)"

"Have you any questions?"

[AFTER QUESTIONS, BEGIN PRACTICE AND ACTUAL TRIALS]

1.11. Expert Observation and Comment-Detailed Information

Three (NTSC) or five (ATV) expert observers are used. After watching the video presentation being tested, the group of experts shall write their comments on paper. To the extent possible one set of comments shall be written that reflect the views of the group. If consensus cannot be found, the observers may write individual comments.

2. DEFINITIONS

2.1. Expert Observer

An expert observer is an individual who has had "recent extensive experience in observing picture quality or impairments, particularly of the type being studied in the subjective test." (CCIR Rec 500-3).

Expert observers shall have normal color vision and visual acuity. Prior to being used as an expert observer, an individual shall have passed a color blindness test (Ishihara or equivalent), a visual acuity test (Snellen Chart or equivalent) that shows the observer has at least 20/20 vision (corrected as necessary) in both eyes, and a contrast sensitivity test (Pelli-Robson or equivalent).

Normally expert observers will be engineers or technicians who work in the television industry. However, for certain impairments an "expert" may be someone who works in the entertainment industry in a creative capacity who would be accustomed to critical inspection of television pictures, e.g., a tape or film editor.

2.2. Non-Expert Observer

A non-expert observer is an individual with no special technical training or experience in television picture quality or impairments.

Non-expert observers shall have normal color vision and visual acuity. Prior to being used as a non-expert observer an individual shall have passed a color blindness test (Ishihara or equivalent), a visual acuity test (Snellen Chart or equivalent) that shows the observer has at least 20/20 vision (corrected as necessary) in both eyes, and a contrast sensitivity test (Pelli-Robson or equivalent).

Normally non-expert observers are regular television viewers.

2.3. D and U Levels

The D and U levels are the power levels of the Desired and Undesired signals, respectively, measured according to procedures described in the Objective and Transmission Test Procedures Manual.

2.4. Unusable

This word is used to describe a television picture that an expert observer would not deem "watchable" by the public.

2.5. Advanced Television System (ATV)

For the purposes of this document, advanced television systems (ATV) refers to any television system whose proponent represents it as having performance superior to that of NTSC.

2.6. Presentation

A presentation is a single picture or motion sequence be it a reference or test picture.

2.7. Test Trial

A test trial is a sequence of reference and test presentation(s) along with an evaluation period.

2.8. Demonstration Trial

A test trial designed for observer orientation. No test results are inferred from demonstration trials.

2.9. Block

A block is a series of contiguous test trials that all deal with a single impairment or with the quality of a single ATV system. There shall be no intermingling of impairments or ATV systems within blocks. Each block shall begin with demonstration trials to familiarize the observers with the test.

2.10. Session

A session is one or more blocks of tests conducted with a single group of observers. No more than thirty to forty minutes of actual testing shall occur within a session without at least a ten minute break.

Within a given session, each block of tests shall involve independent randomizations of the test pictures and the relative orderings of the impairment levels.

3. VIEWING CONDITIONS

3.1. General

All subjective tests shall be conducted in controlled viewing environments. The monitors/display devices shall be calibrated prior to any testing, and the layout and room conditions shall be adhered to throughout the testing.

3.2. Test Viewing Conditions

3.2.1. Viewing Room

The room shall be dimly lit. The exact room illumination shall be such that the requirements below for a dark screen shall be met.

The color temperature of the room lighting shall be D_{6500}.

The room decor shall be neutral and low key. No bright colors shall be used.

Comfortable chairs shall be provided for each observer.

The room shall be quiet and free from aural distractions.

3.2.2. Multiple NTSC Receivers

The 24 consumer NTSC receivers shall be observed from a distance of 5 to 6 picture heights. To the extent possible, the observers shall be ±30° from a perpendicular at the center of the kinescope face.

The setup of these receivers will utilize PLUGE, SMPTE color bar, gray scale, and needle pulse test patterns. Suitable patterns are contained on the laser videodisc, "A Video Standard" (Reference Recordings LD-101).

This videodisc also describes an appropriate procedure for setting proper levels of Brightness, Contrast, Color, and Hue.

After each of the 24 receivers has been set up in accordance with the recorded procedure, it is likely that there will be differences in screen brightness among them. In this event, the brighter receivers shall be reset to match the brightness of the dimmest receiver, while displaying a gray scale test pattern. Matching of, at least, the eight receivers within a given bank shall be the goal if matching of all 24 receivers is not possible.

The peak screen luminance shall be measured and reported.

3.2.3. Large-Screen ATV Display Device

The ATV display device shall be a large screen multi-standard video projector type display.

The 5 observers shall be located at a distance of 3 picture heights from the display along a radius extending from the center of the display screen. They

shall be located within a cone ±30° from a perpendicular extending from the center of the viewing screen.

The wall behind the display device shall be painted a neutral gray and shall be dimly illuminated by D_{6500} light sources. The luminance of these walls shall be 22-38 cd/m^2 (15% of peak luminance of monitor).

The brightness and contrast of the display shall be adjusted with a standard PLUGE signal. The red, blue, and green gains shall be checked with 75% color bars. The overall color temperature of the display shall be adjusted for a peak white color temperature of 6500° K.

The aspect ratio of the display shall be masked for 16:9. The ATV images from all proponents systems shall be displayed at the same height.

At beam cutoff the screen luminance shall not exceed 2% of the peak luminance. The luminance of the screen when displaying black level in a fully dark room shall be about 1% of peak screen luminance.

The following shall be a goal for the ATV display:

The peak luminance of the ATV display shall be adjusted for 150-250 cd/m^2.

3.2.4. Large-Screen NTSC Display Device

NTSC images shall be evaluated at a distance of 5 to 6 picture heights. To the extent possible the observers shall be ± 30° from a perpendicular at the center of the kinescope face.

The display shall be adjusted for a peak screen luminance of 70 ± 10 candelas per square meter. At beam cut-off the screen luminance shall not exceed 2% of peak screen luminance. The luminance of the screen when displaying black level in a fully dark room shall be about 1% of peak luminance.

The wall behind the display device shall be dimly illuminated by D_{6500} light sources. The luminance of the wall shall be 15% of peak screen luminance.

The brightness and contrast of the display shall be set using a standard PLUGE signal. The chroma phase and gain shall be set using SMPTE 75% color bars.

3.3. Use of Receivers and ATV Display

The threshold of visibility and point of unusability, when NTSC is the Desired signal, shall be determined using the 24 NTSC receivers. The steps to be used for NTSC rating tests shall be determined using the large-screen NTSC display. All viewing of ATV signals shall utilize the large-screen ATV display. Identical large-screen displays, both NTSC and ATV, shall be used for expert and non-expert viewing.

4. VIDEO TAPING OF RANGING TESTS

4.1. Threshold Archive Tapes

Video tapes **may** be made showing the impairment at the threshold of visibility. The picture material recorded shall be the same as that viewed by the expert observers during the ranging process. These tapes may not be made at the same time the ranging tests are conducted.

4.2. Recording of Specific Levels of Impairment

The test administrator may, at his discretion, record still and moving pictures of a system under test at various levels of impairment for the purpose of documenting performance at impairment levels which may be of interest to regulatory bodies or to the television community.

4.3. Parent Master Rating Tapes

Parent Master tapes shall be recorded for any future rating tests. These tapes shall contain the selected test pictures impaired at the levels determined during the ranging process. These master tapes shall contain the six levels of impairment which will actually be used to make the rating test tapes. The master tapes also shall contain unimpaired reference pictures at each of the "D" levels under test.

4.4. Making of NTSC Tapes

There is no "typical" NTSC receiver so that any recordings made of NTSC should be regarded as approximately representative of the NTSC receiver population.

Where a recording of NTSC is called for, the following steps shall be used to make the tape:

A. Use the baseband output of an NTSC receiver that is as close to the median (the 12th receiver to show the impairment under study) as possible.

B. Play back this recording and view it on the large-screen NTSC display. Allow the 3 expert observers to compare the large-screen picture with that of the median NTSC receiver. If the two images are comparable, then the tape may be used. If the two are not comparable, then a) another receiver may chosen and the process repeated or b) no NTSC rating tape will made.

It is important to note that not all of the NTSC receivers will have baseband outputs suitable for recording. Also it should be noted that some baseband outputs will have frequency response shaping. Clearly any such receiver with the shaping should not be used.

Most important to note is that is may be impossible at times to make an NTSC recording suitable for later use as it may prove impossible to get a recording that is deemed representative of the median NTSC receiver. If the recording represents the median receiver performance at all impairment levels except the most severe ones, then the expert observers may elect to approve the recording, inasmuch as the lower impairment levels are of the greatest interest.

5. RATING TEST MATERIAL

For the rating tests, the following visual test materials will be used:

Quality: 9 still pictures and 14 motion sequences

Impairment (including Interference): 2 still pictures and 1 motion sequence

Descriptions of these test materials are contained in Appendix 1 of this document.

6. PREPARATION OF RATING VIDEO TAPES

6.1. General Requirements

Each video tape shall have at its beginning the following:

*1 minute of color bars and 400 hz 0 level tone on the designated audio program channel

*30 seconds of a slate containing the following information:

- Copyright Notice
- Serial Number of the Tape
- Name of Proponent System Under Test and Scanning Standard
- Type of Tape: Quality, Parent Master, or Transmission Impairment
- If Parent Master or Transmission Impairment, then the impairment(s) being tested
- Preparation Date of the Tape

* 10 seconds of black

NOTE: If the tape is quality rating, then the above shall be in 1125/59.94 or 1125/60, depending upon the system under test. If the tape is transmission impairment rating, the above shall be in the scanning standard of the system under test.

6.2. Randomization of Test Sequence

The two rating tapes for a given type of test shall use the same test materials, one tape in each of two pseudo-random orders. The two pseudo-randomizations shall be designed to identify, when the results of the tests are analyzed, any effects arising from adjacency of picture material or impairment level or from learning curve. Also as described previously, the same two relative orderings of impairment levels shall be used in Transmission Impairment Rating Tapes for all impairments.

7.1 DATA PRESENTATION

7.1. General

The exact methods to be used to reduce, analyze, and present the data are under study.

After the required number of observers have viewed a particular rating test, the results shall be analyzed and the standard deviation calculated. If the standard deviation is large, the testing laboratory may recommend that additional testing be conducted or that the test method and/or material be modified. It must so advise the Chairman of SS/WP-2, who will determine the best course of action to respond in a timely manner.

7.2. Practice Trials

The first five trials of any rating test shall be deemed practice trials and shall not be used in the analysis of the test results.

8. TEST METHODS FOR INDIVIDUAL IMPAIRMENTS

This section details the particular test method to be used for a given impairment. In some cases an impairment will be first tested by ranging and then by rating so it will appear more than once in the following lists.

8.1. ATV Threshold of Visibility

- Discrete Frequency Interference
- Impulse Noise

8.2. ATV Threshold of Visibility and Point of Unusability

- Multipath
- Airplane Flutter
- UHF Taboos
- Cable ICPM
- Cable Second Order Intermodulation
- Cable Hum and Low-Frequency Noise

8.3. ATV Ranging (3 "D" levels except as noted)

- Random Noise (1 "D" level)
- Co-Channel Interference (2 "D" levels)
- Upper Adjacent Channel Interference (2 "D" levels for Enhanced-NTSC systems)
- Lower Adjacent Channel Interference (2 "D" levels for Enhanced-NTSC systems)
- Cable Third Order Intermodulation (1 "D" level)

8.4. ATV Transmission Impairment Rating (3 "D" levels except as noted)

- Random Noise (1 "D" level)
- Co-Channel Interference (2 "D" levels)
- Upper Adjacent Channel Interference (2 "D" levels for Enhanced-NTSC systems)
- Lower Adjacent Channel Interference (2 "D" levels for Enhanced-NTSC systems)
- Cable Third Order Intermodulation (1 "D" level)

8.5. NTSC Received ATV Threshold of Visibility and Point of Unusability (Enhanced-NTSC systems only)

- Random Noise
- Multipath
- Airplane Flutter
- Impulse Noise (TOV only)

8.6. NTSC Calibration

- Random Noise
- Co-Channel Interference
- Lower-Adjacent Channel Interference
- Upper-Adjacent Channel Interference
- UHF Taboos