RECOMMENDATION ITU-R BT.1438

SUBJECTIVE ASSESSMENT OF STEREOSCOPIC TELEVISION PICTURES

(Question ITU-R 234/11)

(2000)

The ITU Radiocommunication Assembly,

considering

a) that studies are in progress to develop stereoscopic television as a potential future broadcast service;

b) that Recommendation ITU-R BT.1198 has been established for stereoscopic television based on R- and L-eye two channel signals;

c) that subjective assessments are a vital element in the design and introduction of stereoscopic television systems;

d) that shooting conditions, viewing conditions and type of display may influence observer fatigue;

e) that common assessment conditions appropriate for stereoscopic television systems should be established; these conditions should include evaluation methods, shooting conditions, viewing conditions, test materials to be used in the assessment and screening methods to ensure that observers have normal depth perception,

recommends

that the conditions described below should be used for the subjective assessment of stereoscopic television systems.

1 Assessment factors

Assessment factors generally applied to monoscopic television pictures, such as resolution, colour rendition, motion portrayal, overall quality, sharpness, depth, etc., could be applied to stereoscopic television systems. In addition, there would be many factors peculiar to stereoscopic television systems. Some of them are listed below, and further studies are required to identify others and to establish physical definitions.

– *Depth resolution*

  Spatial resolution in depth direction. Coarse resolution in depth direction may reduce picture quality in stereoscopic television.

– *Depth motion*

  A factor related to whether motion or movement along depth direction is reproduced smoothly.

– *Puppet theatre effect*

  This describes one type of distortion in reproduced 3-D images. Stereoscopic objects are sometimes perceived as unnaturally large or small.

– *Cardboard effect*

  This describes another type of distortion in reproduced 3-D images. The 3-D positions of stereoscopic objects are perceived stereoscopically but they appear unnaturally thin.

2 Assessment methods

The methods described in Recommendation ITU-R BT.500 could be applied for the evaluation of the general picture quality of stereoscopic systems as well as sharpness and depth (see Annex 2). When a reference image is available, double-stimulus continuous quality-scale or double-stimulus impairment scale methods can be used. Examples include comparison of display systems, quality assessment of coding systems, and so on. When no reference is available, the categorical judgement method can be used, for example, to identify the merits of stereoscopic systems. Evaluation methods for the assessment of particular factors of stereoscopic television systems require further study.

3 Viewing conditions

Two major factors peculiar to stereoscopic display should be taken into consideration, namely the display frame effect and inconsistency between accommodation and convergence.

Stereoscopic pictures appear highly unnatural when objects positioned in front of the screen approach the screen frame. This unnatural effect is called “the frame effect”. The effect is generally reduced with a larger screen, because observers are less conscious of the existence of the frame when the screen is larger.

The human eye focuses on an object according to the distance to that object. At the same time, we also control the convergence point (gaze point) on the object. Therefore, there is no inconsistency between accommodation and convergence in our everyday life. However when viewing stereoscopic images, the focus point (accommodation) must always be fixed on the screen, independent of the convergence point which is derived from the disparity of the signals. Otherwise, the observer cannot focus clearly. Thus, an inconsistency between accommodation and convergence is introduced in stereoscopic systems.

It is generally said that the minimum value for depth of field of the human eye is ±0.3 D (Diopter: reciprocal value of distance (m)) [Hiruma and Fukuda, 1990]. This means that we can perceive the image without defocusing when the object is located within ±0.3 D. When viewing stereoscopic television, the accommodation point is fixed on the screen, and therefore stereoscopic pictures should preferably be displayed within this range. Since ordinary television programmes include images at infinite distance (that is $D = 0$), the desirable range of depth to be displayed with stereoscopic systems is considered to be within 0 to 0.6 D. Therefore, 0.3 D, i.e. 3.3 m, is considered to be the optimum viewing distance.

Camera parameters (camera separation, camera convergence angle, focal length of lens), resolution of the system and the frame effect should be taken into account in determining viewing conditions (screen size). In the case of HDTV when watching at the standard viewing distance of 3 *H* (*H* denotes picture height), the viewing distance of 3.3 m corresponds to a 90-inch screen. In the case of standard definition television (SDTV) when watching at the standard viewing distance of 6 *H*, this distance corresponds to a 36-inch screen. A subjective assessment of the relationship between screen size and depth perception was carried out with stereoscopic HDTV system, and the results showed that the most natural depth perception was obtained with a screen size of 120 inches, which corresponds to viewing distance of 2.2 *H* [Yamanoue *et al.*, 1997].

4 Observers

Observers should have normal acuity (see Recommendation ITU-R BT.500). In addition, they should have normal stereopsis. In order to check their stereopsis, vision test materials listed in Annex 1 can be used.

5 Test materials

Test materials for screening observers and still and motion sequences of natural scenes are listed in Annex. 1.

The 3-D effects obtained from stereoscopic pictures depend largely on the shooting conditions, such as camera separation, camera convergence angle and focal length of the lens. The motion sequences were shot under the condition of camera separation of 65 mm, corresponding to average eye separation, and most of them were produced with the uncrossed camera layout, which gives an orthostereoscopic condition [Yamanoue *et al.*, 1998].

REFERENCES

HIRUMA, N. and FUKUDA, T. [December, 1990] Accommodation response to binocular stereoscopic TV images and their viewing conditions. *J. SMPTE*, 102, **12**, p. 2047-2054.

YAMANOUE, H. *et al.* [October, 1997] Subjective study on the orthostereoscopic conditions for 3D-HDTV. ITE Tech. Report, Vol. 21, **63**, p. 7-12.

YAMANOUE, H. *et al.* [1998] Orthostereoscopic conditions for 3-D HDTV. *Proc. SPIE*, 3295, *Stereoscopic displays and Applications IV*.

ANNEX 1

Test materials for subjective assessment of stereoscopic television pictures

1 Vision test

Table 2 lists the test charts for the vision test. These 12 tests are selected according to the hierarchy of the human visual system from lower to higher levels. Eight main vision tests (VTs) are described below, and the other four are for the clinical test. Observers must have normal stereopsis, meaning that they must pass test VT-04 for fine stereopsis and test VT-07 for dynamic stereopsis. The remaining six tests are for more detailed characterization. The test charts should be viewed from three times the height of the display screen (3 *H*).

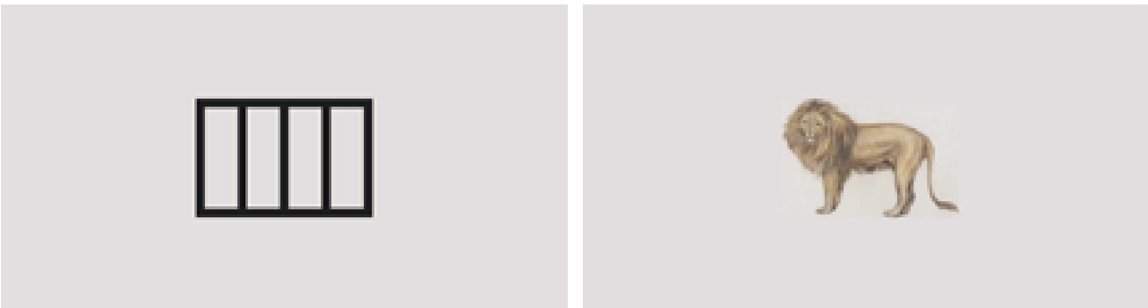
Below, right and left thumbnail images are put side by side for crossed free fusion for explanatory purposes.

a) *VT-01*: Simultaneous perception (lion test)

Tests the ability to perceive dichoptically presented images simultaneously and in the correct position. A cage image is presented to one eye and a lion image to the other eye, with its position moving by 12′/s. The size of each image is fixed at 10° so that the observers can capture the images within their paramacula. Observers with normal vision can see the lion in the cage at a certain time within the presentation period.

FIGURE 1

Test chart for VT-01



Right image Left image

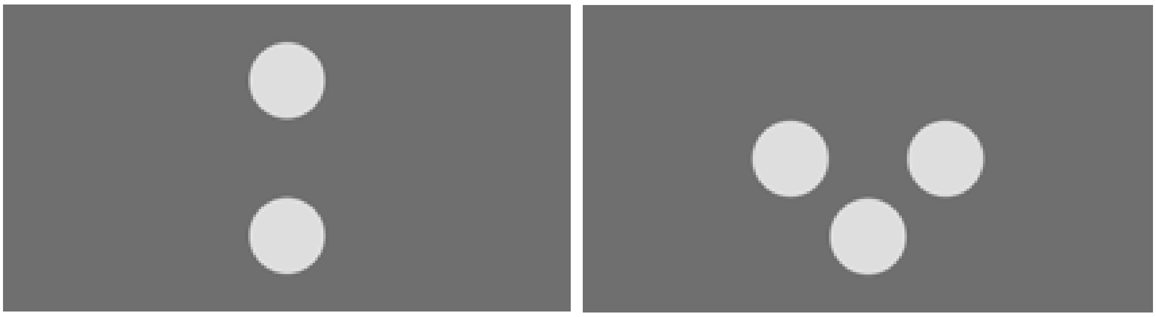
1438-01

b) *VT-02*: Binocular fusion (worth 4-dot test)

Tests the ability to perceive two dichoptic images in left and right eyes as one image. The image for one eye has two dots, and the image for the other eye has three dots, with one dot in common. Observers with normal vision can see four dots.

FIGURE 2

Test chart for VT-02



Right image Left image

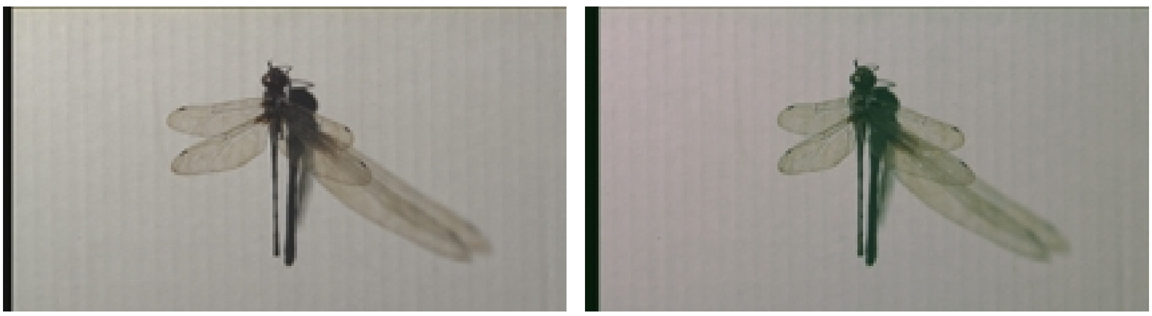
1438-02

c) *VT-03*: Coarse stereopsis (dragonfly test)

Tests the ability to perceive dichoptically presented images with a parallax as one image with a coarse depth. The images for the two eyes are a stereopair of images of a dragonfly with its wings spreading. Observers with normal vision can perceive the wings in front of the display screen.

FIGURE 3

Test chart for VT-03



Right image Left image

1438-03

d) *VT-04*: Fine stereopsis (circle test)

Tests the ability to perceive dichoptically presented images with a parallax as one image with a fine depth. Nine test lozenge patches are provided and each of them has four circles in which only one circle has a small parallax. Observers with normal vision can perceive the circle with a small parallax in front of the display screen. Table 1 shows the test number, correct answers, and angle of stereopsis at 3 *H*.

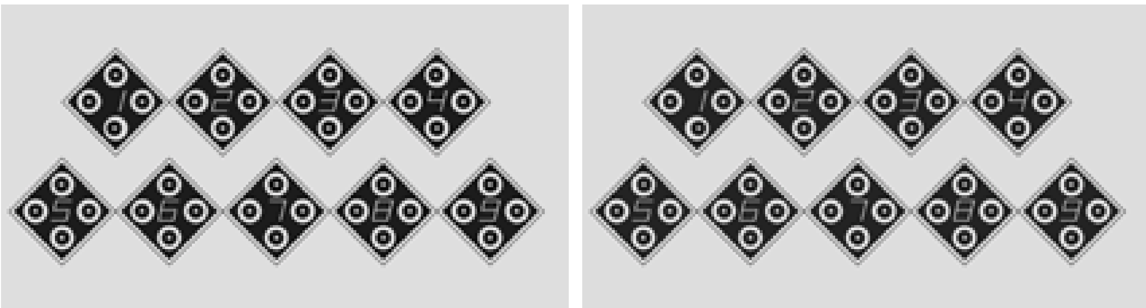
TABLE 1

Correct answers and parallax

Test No.	Correct answers	Angle of stereopsis at 3 *H* (")
1	Bottom	480
2	Left	420
3	Bottom	360
4	Top	300
5	Top	240
6	Left	180
7	Right	120
8	Left	60
9	–	0

FIGURE 4

Test chart for VT-04



Right image Left image

1438-04

e) *VT-05*: Crossed fusional limit (bar test)

Tests the ability to perceive dichoptically presented images with crossed disparities as one image. A stereopair of bars is presented with its parallax changing by 10′/s. The fusional limits for the ascending and the descending series can be measured. Observers are instructed to report their fusional break as soon as they perceive double images in the ascending series, and their recovery of fusion as soon as they perceive the dichoptic images as a single image in the descending series.

FIGURE 5

Test chart for VT-05

Right image Left image

1438-05

f) *VT-06*: Uncrossed fusional limit (bar test)

Tests the ability to perceive dichoptically presented images with uncrossed disparities as one image. Presented images are the same as in the crossed case above, but right and left images are swapped.

FIGURE 6

Test chart for VT-06

Right image Left image

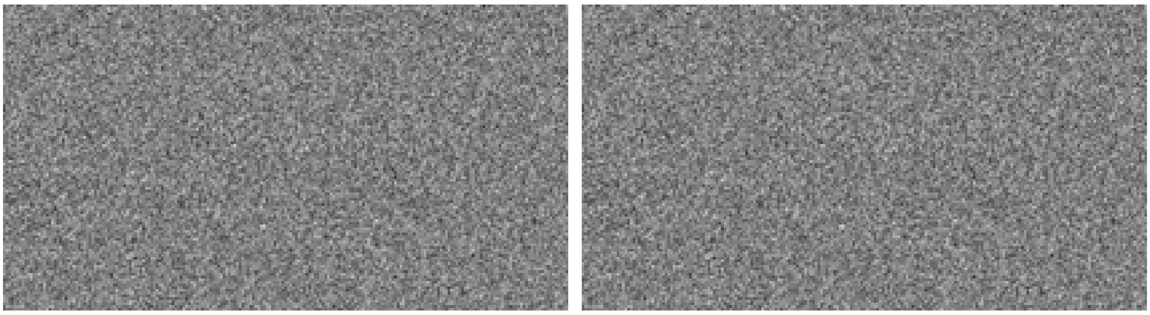
1438-06

g) *VT-07*: Dynamic stereopsis (dynamic random dot stereogram test)

Tests the ability to perceive depth in moving random dot stereogram images. Observers with normal vision can perceive a rectangular shape and a sinsoidal depth motion in the dynamic random dot stereogram.

FIGURE 7

Test chart for VT-07



Right image Left image

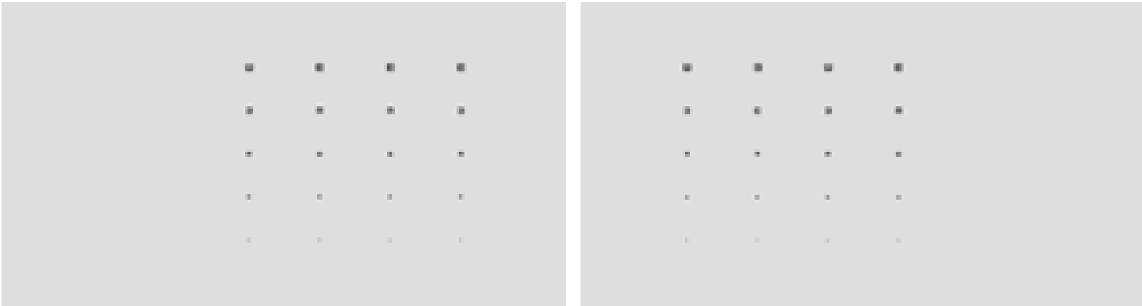
1438-07

h) *VT-08*: Binocular acuity (acuity test)

Tests the binocular acuity with binocular fusion, including any imbalance of monocular acuity which might prevent good stereopsis. The images have four columns and five lines which consists of E characters with a variety of orientation and size. The centre two columns can be seen with both eyes; the left two columns can be seen only with the left eye; and the right two columns can be seen only with the right eye. Observers with normal vision can tell the orientation of the E characters correctly. The character sizes correspond to acuities of about 1.0, 0.5, 0.33, 0.25, and 0.125 at 3 *H*.

FIGURE 8

Test chart for VT-08



Right image Left image

1438-08

2 Natural images

The natural images consist of 15 still pictures and 15 motion sequences, as listed in Tables 3 and 4. Some of them are illustrated in Appendix 1. Each image is printed from left to right as left image, right image, left image: the 3-D image can be obtained by fusing the leftmost pair (eyes uncrossed) or the rightmost pair (eyes crossed).

3 Status of usage of the stereoscopic test material

Use of the test material is restricted to the following purposes only:

- Technical evaluation, including:
  - research and development of equipment and systems,
  - testing of equipment in development and production process,
  - testing of transmission condition for broadcasting and telecommunication,
  - maintenance of equipment.
- Demonstration, including:
  - presentation at technical conference and workshop,
  - presentation of performance and functionality of equipment, excluding commercial promotion.

NOTE 1 – Presentation of the motion sequence No. 10 – Football, is allowed ONLY at research facilities such as universities, research institutes, and manufacturer's laboratories, not in public.

TABLE 2

Stereoscopic test materials – VT

No.	Item	Test for	Content
1	Simultaneous perception	The ability to perceive dichoptically presented images simultaneously and in the correct position	A cage image is presented to one eye and a lion image to the other eye
2	Binocular fusion	The ability to perceive two dichoptic images in left and right eyes as one image	The image for one eye has two dots, and the image for the other eye has three dots, with one dot in common
3	Coarse stereopsis	The ability to perceive dichoptically presented images with a parallax as one image with a coarse depth	The image for two eyes are a stereopair of images of a dragonfly with its wings spreading
4	Fine stereopsis	The ability to perceive dichoptically presented images with a parallax as one image with a fine depth	Nine test lozenge patches are provided and each of them has four circles in which one circle has a small parallax
5	Crossed fusion limit	The ability to perceive dichoptically presented images with crossed disparities as one image	A stereopair of bars is presented with its crossed parallax changing by 10′/s
6	Uncrossed fusional limit	The ability to perceive dichoptically presented images with uncrossed disparities as one image	A stereopair of bars is presented with its uncrossed parallax changing by 11′/s
7	Dynamic stereopsis	The ability to perceive depth in moving random dot stereogram images	Dynamic random dot stereogram
8	Binocular acuity	The binocular acuity, including any imbalance of monocular acuity which might prevent good stereopsis	E characters with a variety of orientation and size
9	Horizontal strabismus	The horizontal deviation of the eye which the patient cannot overcome	Vertical and horizontal lines
10	Vertical strabismus	The vertical deviation of the eye which the patient cannot overcome	Vertical and horizontal lines
11	Aniseikonia	A condition in which the ocular image of an object as seen by one eye differs in size and shape from that seen by the other	The left image consists of the characters "[o" and the right consists of the characters "o]" where the "o" character position is common
12	Cyclophoria	The deviation of one eye or the other around the anteroposterior axis when fusion is prevented	The left image consists of the face of a clock and the right consists of the hands of a clock at six o'clock

NOTE 1 – These materials are recorded on digital video tape recorder (VTR) in the 1125/60/2:1 format (see Recommendation ITU-R BT.709).

NOTE 2 – The materials can be obtained from the Institute of Image Information and Television Engineers (ITE), 3-5-8 Shibakoen, Minato-ku, Tokyo 105-0011, Japan, Phone: +81-3-3432-4677, Fax: +81-3-3432-4675, e-mail: ite@ite.or.jp.

TABLE 3

Stereoscopic test materials – still pictures

No.	Title	Content	Representative of	Major factors to be assessed	Keystone distortion
1	Autumn tints	Red autumnal *momiji* leaves against the light	Outdoor shooting	Still and depth resolution	No
2	Autumn tints and Buddhist temple	A scene with pure red *momiji* leaves in direct light, and a Buddhist temple in the background	Outdoor shooting	Still and depth resolution	No
3	Attractive Japanese kimono in a Buddhist temple	Woman dressed in kimono with *Daikakuji* temple in the background	Outdoor shooting	Still and depth resolution	No
4	Autumn leaves	Woman dressed in kimono in a Japanese garden filled with autumn leaves	Outdoor shooting	Still and depth resolution	No
5	Sky	Scenery of trees with leaves in autumn tints	Outdoor shooting	Still and depth resolution	No
6	Under the shade of a tree	Woman in a grove dressed in kimono	Outdoor shooting	Still and depth resolution	No
7	By the side of the autumn tints	Woman dressed in kimono and autumn tints in the temple grounds	Outdoor shooting	Still and depth resolution	No
8	Japanese garden	Autumn tinted *Eikando* garden	Outdoor shooting	Still and depth resolution	No
9	Beauty in kimono	Autumn tints and a lady dressed in kimono	Outdoor shooting	Still and depth resolution	No
10	City scenery 1	Modern building and woman	Outdoor shooting	Still and depth resolution	No
11	City scenery 2	Artificial waterfall and woman	Outdoor shooting	Still and depth resolution	No
12	City scenery 3	A building promenade and woman	Outdoor shooting	Still and depth resolution	No
13	In my room 1	Woman at ease in her room	Studio production	Still and depth resolution	No
14	In my room 2	Woman at ease in her room	Studio production	Still and depth resolution	No
15	Dining	Scenery of a woman at a dining table	Studio production	Still and depth resolution	No

NOTE 1 – These materials are recorded on digital VTR in the 1125/60/2:1 format (see Recommendation ITU-R BT.709).

NOTE 2 – The following shooting conditions apply for all the materials: lens f = 75 mm, film EPR70 mm, camera separation 60 mm, with an uncrossed camera layout.

NOTE 3 – The materials can be obtained from the Institute of Image Information and Television Engineers (ITE), 3-5-8 Shibakoen, Minato-ku, Tokyo 105-0011, Japan, Phone: +81-3-3432-4677, Fax: +81-3-3432-4675, e-mail: ite@ite.or.jp.

TABLE 4

Stereoscopic test materials – motion sequences

No.	Title	Content	Representative of	Major factors to be assessed	Motion	Focal length (mm)	Keystone distortion
1	Tulip garden	A girl walking in a tulip garden	Outdoor shooting	Still and depth resolution	Slow	40	No*
2	Festival	A portable shrine and paper storm	Outdoor shooting	Still and depth resolution	Slow	12	No*
3	Portable shrines	Portable shrines carried out	Outdoor shooting	Still and depth resolution	Slow pan	20	No*
4	Crossing ships	Crossing ships and audience	Outdoor shooting	Depth motion	Slow pan	40	No*
5	Red leaves	A woman and red leaves	Outdoor shooting	Depth motion	Medium	12	No*
6	Botanical gardens	Waterfall in botanical gardens	Outdoor shooting	Orthostereoscopy	Fixed	12	No*
7	Living room	A woman sitting on a sofa	Studio production	Orthostereoscopy	Fixed	12	No*
8	A meal	People around a meal on a table	Studio production	Orthostereoscopy	Fixed	12	No*
9	Amusement park	Girls playing in an amusement park	Outdoor shooting	Depth motion	Medium	12	No*
10	Football	A football game	Outdoor shooting	Depth motion and resolution	Medium	12	No*
11	A vocalist	A vocalist in a hall	Studio production	Orthostereoscopy	Fixed	12	No*
12	Chromakey	A woman and flowers	Studio production	Chromakey	Fixed	12	No*
13	Flower pot	A girl and flower pot	Outdoor shooting	Depth motion	Medium	12	Yes
14	An aquarium	Tropical fish in an aquarium	Outdoor shooting	Depth motion	Fixed	12	Yes
15	Flower garden	A girl walking in a flower garden	Outdoor shooting	Depth motion and resolution	Slow	12	Yes

NOTE 1 – These materials are recorded on digital VTR in the 1125/60/2:1 format (see Recommendation ITU-R BT.709).

NOTE 2 – The materials with * are produced with an uncrossed camera layout.

NOTE 3 – Camera separation is 65 mm for all the materials.

NOTE 4 – The materials can be obtained from the Institute of Image Information and Television Engineers (ITE), 3-5-8 Shibakoen, Minato-ku, Tokyo 105-0011, Japan. Phone: +81-3-3432-4677, Fax: +81-3-3432-4675, e-mail: ite@ite.or.jp.

APPENDIX 1

TO ANNEX 1

Examples of natural motion sequences

FIGURE 9

No. 1 - Tulip garden

Left image Right image Left image

1438-09

FIGURE 10

No. 2 - Festival

Left image Right image Left image

1438-10

FIGURE 11

No. 5 - Red leaves

Left image Right image Left image

1438-11

FIGURE 12

No. 7 - Living room

Left image Right image Left image

1438-12

ANNEX 2

Experimental results using the double-stimulus continuous quality scale (DSCQS) method

This Annex documents the application of the DSCQS method to the subjective evaluation of such pictures. The DSCQS method has been used widely and successfully to evaluate the subjective image-quality of monoscopic pictures. Experiences have shown that this method is equally well-suited to evaluating stereoscopic pictures, and that it can easily be adapted to measure other image attributes in addition to image-quality, such as perceived sharpness and depth.

1 Measuring perceived sharpness and depth with the DSCQS method

The DSCQS method was easily adapted to measure attributes other than subjective image-quality. This was done by making specific changes to the instructions that were given to viewers. For example, the method was adapted to measure perceived sharpness, and the overall impression of depth of stereoscopic image sequences. Only one attribute (either subjective image-quality, perceived sharpness or perceived depth) was measured in any given testing session.

2 Illustrative study using the DSCQS method with stereoscopic images

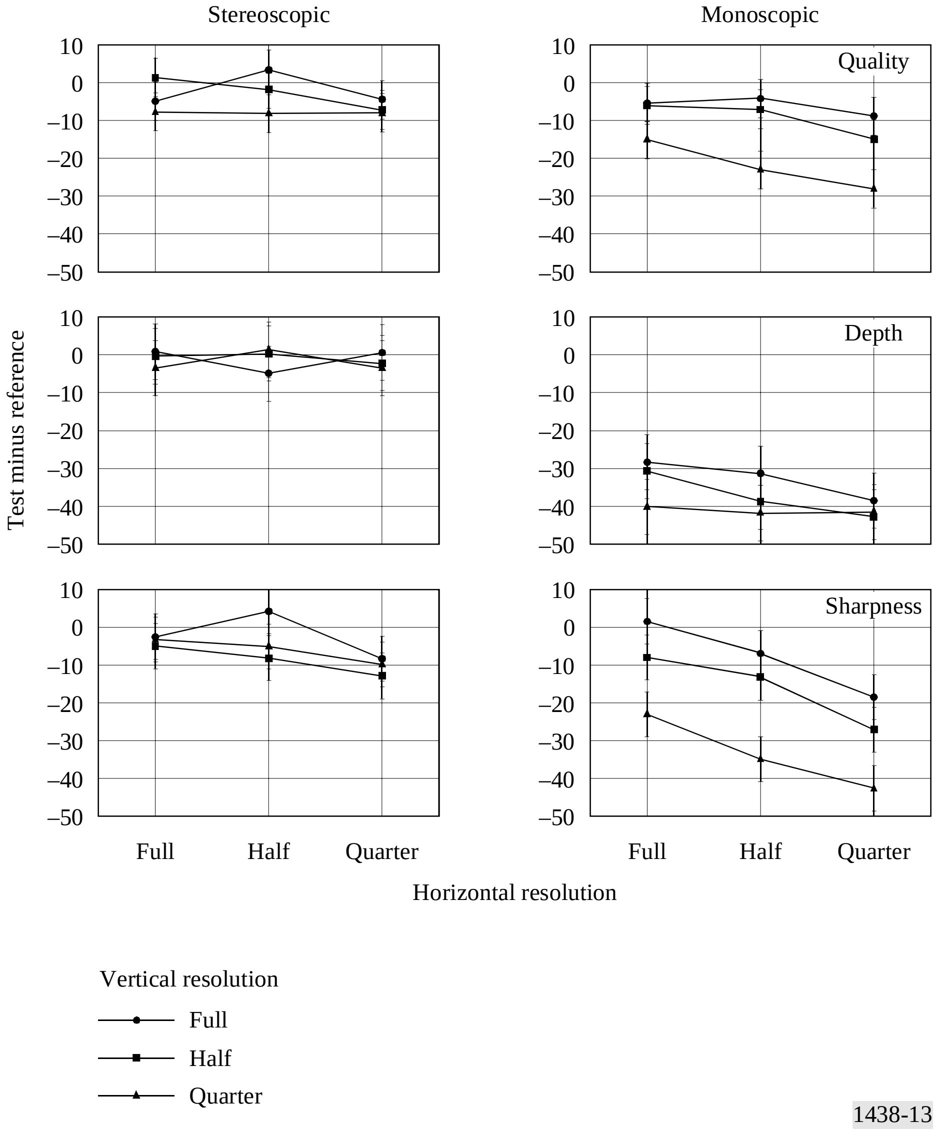
The DSCQS method was used to measure the subjective quality, perceived sharpness, and overall depth of a set of stereoscopic and monoscopic image sequences. In the illustrative study, the goal was to determine whether processing one channel of a stereoscopic image sequence would affect these attributes. To this end, the right-eye view of the stereoscopic image sequences was low-pass filtered at three levels: unfiltered, half-resolution and quarter-resolution. In the monoscopic conditions, both eyes saw the filtered view. A survey of the scientific literature [Julesz, 1971; Pastoor, 1991; Pastoor *et al*., 1995; Perkins, 1992 and Berthold, 1997] led us to expect that filtering one channel of a stereoscopic image would have a much smaller effect on subjective ratings than filtering both channels – subjective rating would be dominated by the unfiltered channel.

In the set-up, the left and right image pairs of a stereoscopic sequence were displayed at 120 Hz using a time-sequential method, on a 29 inch direct-view video monitor or on a 65 inch rear projector. Images for the left and the right eyes were interleaved temporally, and displayed in synchrony with the opening and closing of a pair of Crystal Eyes liquid-crystal shutter-glasses, manufactured by StereoGraphics. The liquid-crystal glasses had a light transmittance of about 30%, and response times of 0.2 ms and 2.8 ms for closing and opening, respectively. This meant that the peak luminance entering the observers' eyes was about 21 cd/m^2, dimmer than desirable, but the best achievable with the available time-sequential display technology. Viewing distance was 4 *H*. A gray field of 10 cd/m^2 was displayed between presentations of the A and the B sequences. It should be noted, that any other method of displaying stereoscopic images could readily be substituted for the time-sequential method without affecting the use of the DSCQS method.

Results of the experiment are presented in Fig. 13. The Y-axis indicates the test minus reference score. A score of zero means that the test sequence was rated equivalently to the unprocessed stereo reference sequence. A negative score means that the test sequence was rated lower than the reference sequence.

FIGURE 13

Results of the illustrative study using the DSCQS method



Effects of low-pass filtering are evident in the slope and vertical offset of the lines. As expected, in monoscopic conditions (see right panels of Fig. 13), low-pass filtering had a large effect on rated sharpness and image quality. The low ratings of depth occurred because only monocular cues to depth were available in the monoscopic test sequences. Also as expected, in the stereo conditions, (see left panels of Fig. 13) all three dimensions (quality, depth and sharpness) were rated higher than in the monoscopic conditions. Low-pass filtering of one channel of a stereo pair caused negligible effects on perceived depth, and minor effects on perceived sharpness and overall quality. Evidently, the high spatial frequency information in the unfiltered left eye image compensated for the lack of this information in the right eye image.

The illustrative study, and other work on stereoscopic image sequences with the DSCQS method [Stelmach and Tam, 1998] leads to a conclusion that this method is a valuable and useful tool for studying stereoscopic pictures. The method could be adapted to measure other aspects of stereoscopic image sequences such as presence, power, and naturalness.

REFERENCES

BERTHOLD, A. [1997] The influence of blur on the perceived quality and sensation of depth of 2D and stereo images. *ATR Human Information Processing Research Laboratories Technical Report*, TR-H-232, Kyoto, Japan.

JULESZ, B. [1971] Foundations of Cyclopean Perception. *The University of Chicago Press*, Chicago, IL United States of America.

PASTOOR, S. [1991] 3D-television: A survey of recent research results on subjective requirements, *Signal Processing: Image Communication*, 4(**1**), p. 21-32.

PASTOOR, S., WÖPKING, M., FOURNIER, J. and ALPERT, T. [1995] Digital stereoscopic imaging & applications (DISTIMA): Human Factors Data, Deliverable ID: R2045/HHI/AT/DS/C/026/b1.

PERKINS, M. G. [1992] Data compression of stereopairs. *IEEE Trans. on Comm*., 40(**4**), p. 684-696.

STELMACH, L. and TAM, W. J. [1998] Stereoscopic image coding: effect of disparate image-quality in left- and right-eye views, *Signal Processing: Image Communication,* **14**, p. 111-117.