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*).

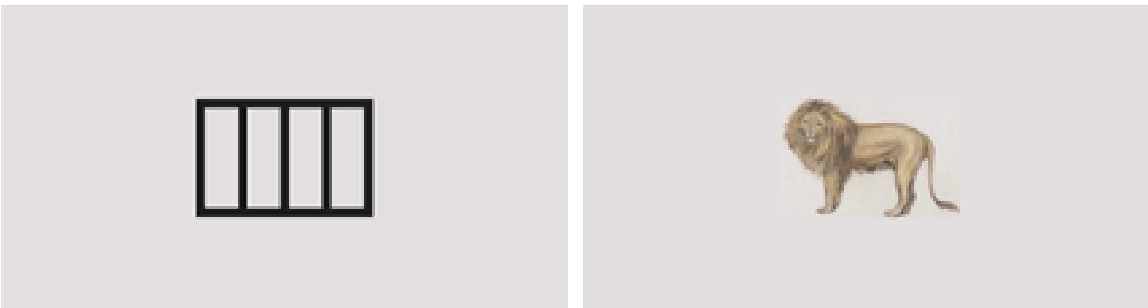
Below, right and left thumbnail images are put side by side for crossed free fusion for explanatory purposes.

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