I n t e r n a t i o n a l   T e l e c o m m u n i c a t i o n   U n i o n

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.916

(03/2016)

SERIES P: TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

Audiovisual quality in multimedia services

Information and guidelines for assessing and minimizing visual discomfort and visual fatigue from 3D video

Recommendation ITU-T P.916

ITU-T

International
Telecommunication
Union

ITU-T P-SERIES RECOMMENDATIONS

TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
Voice terminal characteristics	Series	P.30 P.300
Reference systems	Series	P.40
Objective measuring apparatus	Series	P.50 P.500
Objective electro-acoustical measurements	Series	P.60
Measurements related to speech loudness	Series	P.70
Methods for objective and subjective assessment of speech quality	Series	P.80 P.800
Audiovisual quality in multimedia services	**Series**	**P.900**
Transmission performance and QoS aspects of IP end-points	Series	P.1000
Communications involving vehicles	Series	P.1100
Models and tools for quality assessment of streamed media	Series	P.1200
Telemeeting assessment	Series	P.1300
Statistical analysis, evaluation and reporting guidelines of quality measurements	Series	P.1400
Methods for objective and subjective assessment of quality of services other than voice services	Series	P.1500

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T P.916

Information and guidelines for assessing and minimizing visual discomfort and visual fatigue from 3D video

Summary

Recommendation ITU-T P.916 addresses issues related to assessing and minimizing visual discomfort and visual fatigue from 3D video. Issues addressed include 3D video characteristics that cause visual discomfort; symptoms of visual fatigue; relationships among the camera positions, the captured content and the viewer's perception; footage guidelines; viewer guidelines; information on the 3D; and viewing environment considerations.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.916	2016-03-15	9	11.1002/1000/12778

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Introduction

Three dimensional television (3DTV) systems are based on binocular disparity. Compared to the real world scenario, the reproduction of depth in 3DTV systems is limited by the camera position, number of cameras, source content and the technical limitations of the display technique. These include resolution, frame rate and view separation conditions. Current display technologies may also induce a conflict between accommodation (eyes focusing on the screen) and the convergence (eyes converging to a point in front of or behind the screen). These limitations can cause visual discomfort and visual fatigue.

Recommendation ITU-T P.916

Information and guidelines for assessing and minimizing visual discomfort and visual fatigue from 3D video

1 Scope

This Recommendation includes information and guidelines for assessing and avoiding visual discomfort and visual fatigue from 3D video. It describes potential causes of visual discomfort and symptoms of 3D visual fatigue, including problems caused by watching stereoscopic and autostereoscopic television. This information is primarily intended to be used for the design and conduct of subjective assessment of the video or audiovisual quality of 3D video. This information is also appropriate for all 3D viewing experiences.

1.1 Limitations

This Recommendation includes guidance on recognizing visual fatigue symptoms and identifying 3D programming clip characteristics that may cause visual fatigue.

This Recommendation excludes objective assessment methods of 3D visual fatigue, because these are not computational algorithms but rather correspond to a medical diagnostic process performed by medical doctors.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T P.913] Recommendation ITU-T P.913 (2016), *Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment*.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 accommodation-vergence conflict: Accommodation (eyes focussing on an object) and convergence (eyes pointing toward the object) are normally linked. Accommodation-vergence conflict is a conflict between accommodation (eyes focusing on the screen) and the convergence (eyes converging to a point in front of or behind the screen). Current 3D display technologies may induce accommodation-vergence conflict.

3.2.2 binocular disparity: The two eyes view an object from two different positions and angels. Due to the paralax, the object is projected at different positions on the retina of each eye. The differences of the position between these two projections is the binocular disparity. This can be the strongest depth cue for close objects.

3.2.3 binocular depth cues: Binocular depth cues stem from binocular disparity and the state of the oculomotor system.

3.2.4 cardboard effect: The cardboard effect occurs when the imaging and display conditions reduce the reproduction magnification ratio of depth directions and distort the perception of objects with visually imperceptible thickness. The 3-D positions of stereoscopic objects are perceived stereoscopically but they appear unnaturally thin.

3.2.5 depth cues: Visual indicators of the relative distances between the observer and viewed objects. There are binocular and monocular depth cues.

3.2.6 depth distortion: Depth distortion is the discrepancy between the displayed depth reconstruction and real world. The cardboard effect is one example of a depth distortion.

3.2.7 depth motion: Objects moving in depth (e.g., toward the viewer).

3.2.8 depth resolution: spatial resolution in depth direction.

3.2.9 frame effect: 3D 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.

3.2.10 medical signs: Medical signs are physiological reactions that are measured and interpreted within a controlled procedure.

3.2.11 medical symptoms: Medical symptoms are expressed by the observer and constitute an indication of his or her perceived mental or physical state.

3.2.12 planar motion: Motion with constant depth (i.e., within a plane perpendicular to the observer).

3.2.13 puppet theatre effect: This is an example of size distortion. The reproduction magnification ratio of an object at the shooting distance (the perceived size) varies with the imaging and display conditions. The resulting distortion in size may make an object be perceived as unnaturally small.

3.2.14 size distortion: Distortions in the 3D geometry may occur due to inconsistencies between the capturing situation (notably the camera settings) and the reproduction (notably the display settings). The puppet theatre effect is an example.

3.2.15 monocular depth cues: Monocular depth cues can be perceived from a single eye's view. They include occlusion (objects hidden behind other objects), relative sizes of known objects, vanishing point perception, blur from the focus, motion parallax, light, shade, texture gradient, aerial perspective, height in the visual field, kinetic depth and others.

3.2.16 visual comfort zone: Visual comfort zone is defined as the depth interval which allows for 3D viewing without the introduction of visual discomfort.

3.2.17 visual discomfort: Visual discomfort is a negative sensation that is subjectively reported. Visual discomfort may be triggered by the perception of 3D presentations. Visual discomfort is usually a transient state that disappears quickly after the person stops watching the problematic stimuli. The amount and duration of visual discomfort is highly individual. Visual discomfort may have other causes, such as visual flicker.

3.2.18 visual fatigue: Visual fatigue is generally caused by the repetition of excessive visual effort and can build up as these efforts are repeated. It disappears after sufficient rest. Visual fatigue is a medical state that may be objectively diagnosed by the presence of objective indications (medical signs) and subjective indications (medical symptoms). Visual fatigue can only be diagnosed by a medical doctor.

3.2.19 window violation: Window violation occurs when an object that is in front of the display overlaps the border of the screen. The 3D screen is often perceived as a window and in the real world objects can only be cropped when they lie behind the window. However, objects with crossed disparity are also cropped. When this occurs, the part of the object within the screen appears to be in front of the screen, while the cropping conflicts with this perception, by implying that the object should be at the depth of the screen or behind the screen.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

3DTV Three Dimensional Television

ABV Abnormal Binocular Vision

HD High Definition

HRC Hypothetical Reference Circuit

HUD Head-up Display

PC Pair Comparison

SD Standard Definition

5 Conventions

"3D" in this Recommendation refers to a technology that projects dedicated views for each eye. The ultimate goal is that these views depict the same situation as would be seen in reality (e.g., horizontally shifted). The technology as a whole can include more than two views, as per a multi-view autostereoscopic display. This can be noted as S3D for stereoscopic TV, or 3DMV for multiview.

6 Technological issues for 3D displaying

This clause describes background information required to understand technological issues for displaying 3D content, including constraints of camera capture, transmission, display and the human visual system.

6.1 Accommodation-vergence conflict

Accommodation (eyes focussing on an object) and convergence (eyes pointing toward the object) are normally linked. Current 3D display technologies cause a conflict because accommodation and convergence do not follow the usual pattern. With 3D displays, the eyes focus on the screen but the eyes converge to a point in front of or behind the screen. This is the accommodation-vergence conflict.

The accommodation-vergence conflict has a static component (as described above) and also has a temporal component due to a link within the vision system. The accommodation control system and the vergence ocular motor system are linked. Changes in accommodation automatically induce changes in vergence and vice versa. This is the temporal component of the accommodation-vergence conflict. This is in conflict with current 3D display technology, because the accommodation always stays on the screen while the vergence changes according to the virtual object's position (binocular disparity).

More research is needed, because [b-Shiomi] indicates that there is no accommodation-vergence conflict.

6.2 3D geometric distortions

A disagreement between monocular and binocular depth cues may lead to visual discomfort. Disparity is the most important binocolular depth cue.

In case of multi-view transmission, including stereoscopic transmissions, the resolution of the image impacts on the precision of the horizontal disparity and thus of the depth resolution. Likewise, in multi-view plus depth representations, the resolution of the depth map limits the picture quality of synthesized views.

The quality of the depth rendering is affected by both recording and display conditions which are closely linked.

In order to respect the 3D geometry of the scene during the reconstitution, it is necessary to have an agreement between the different depths cues (monocular and binocular) involved in the scenes. For example in the case of a cardboard effect, the monocular depth cues describe a continuous variation of depth but the binocular depth cues show less variation of depth than the monocular ones. These contradictions between the different depth cues are something unnatural and the scene looks layered.

To ensure respect of 3D geometry, it is then necessary to consider the characteristics of the video camera and the target display at the moment of the capture to ensure that there are no conflicts between monocular and binocular depth cues [b-Chen2] [b-Woods]. The distortion of the depth rendering can be evaluated as follows:

$$D_s = \frac{Vb}{Bz+Mfb \times \left(1-\frac{z}{d_{cov}}\right)}$$

With:

f:	the focal length of the camera
b:	the inter-camera baseline (distance between focal length)
w_ccd,*h_ccd*:	camera sensor width and height
d_cov:	the convergence distance
D_s:	depth distortion
(*x*,*y*,*z*):	the position of a point in camera space (in front of the camera)
B:	inter-pupil baseline: the distance between the eyes of the observers
*W*_screen:	the width of the display
M:	the magnification factor: *M*=*W*_screen/*w_ccd*

These factors should be taken into account during capture. If not, conflict between depth cues will be revealed and distortion of the 3D geometry of objects will appear.

Figure 1 and Figure 2 provide examples of different capture and display settings with their rendering on a specific display [b-Woods].

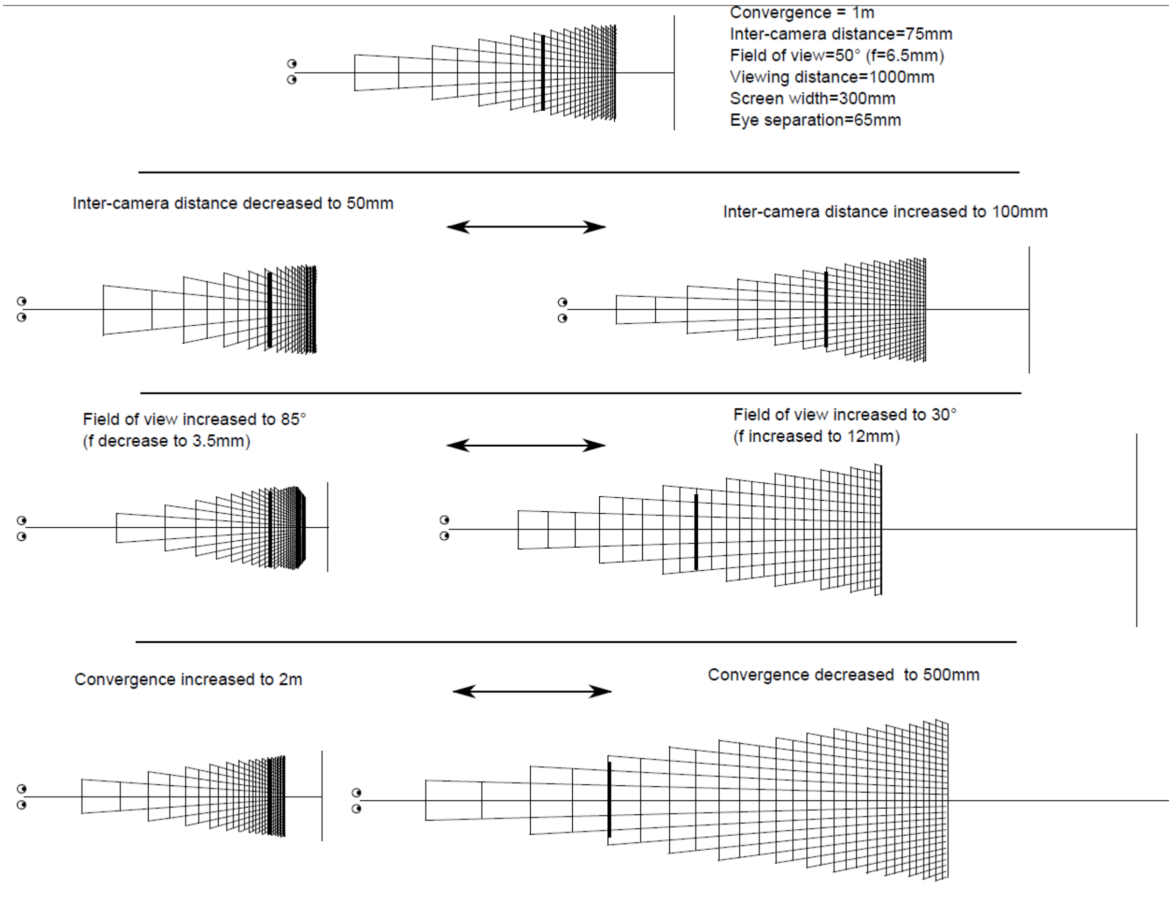


Figure 1 – Examples of different capture settings and their respective effects on the depth rendering on a fixed display [b-Woods]

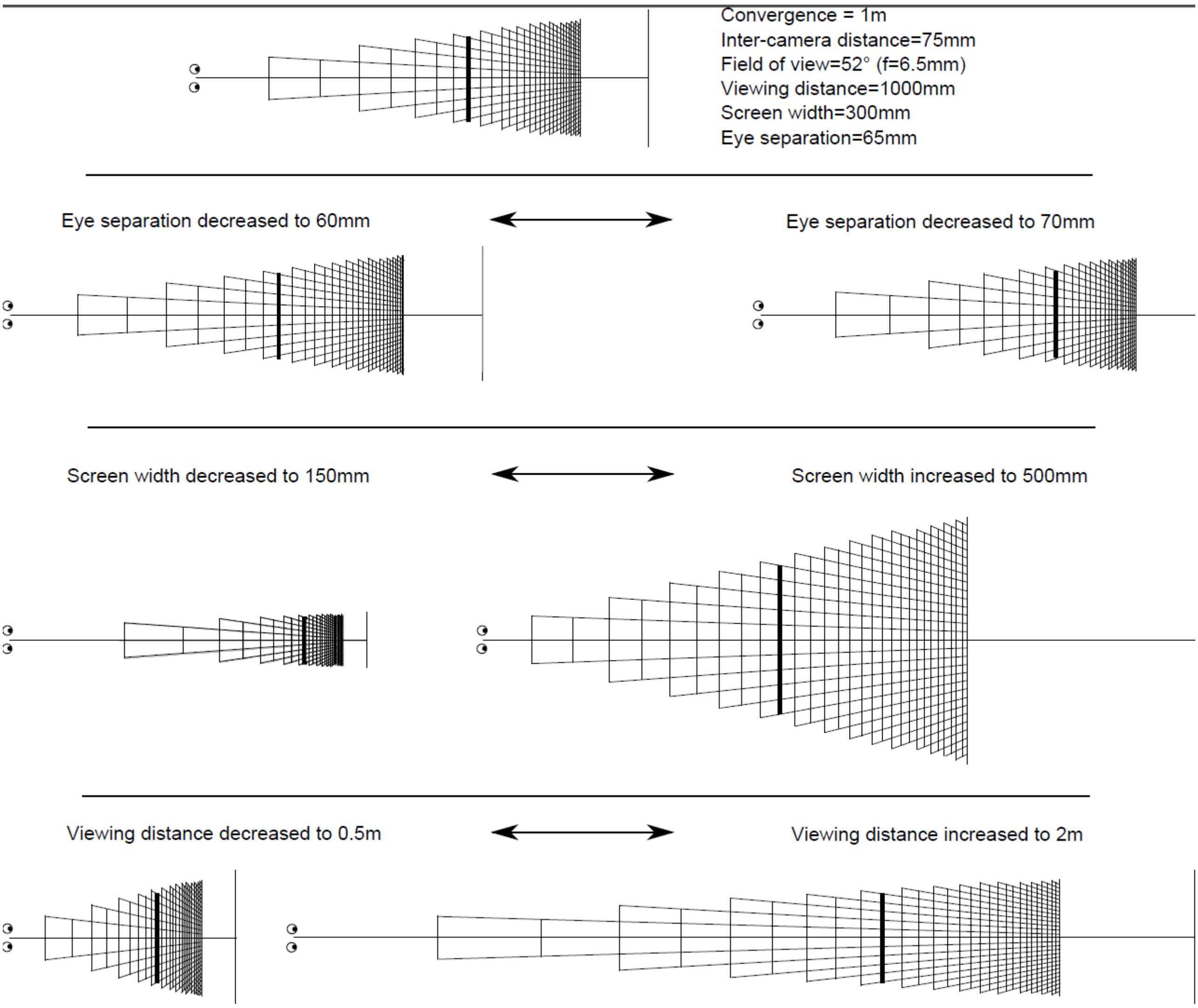


Figure 2 – Examples of different rendering settings and their respective effects on the depth rendering of content captured with fixed settings[b-Woods].

6.3 Visual comfort zone or comfortable viewing zone

See clause 3.2 for the definition of the term 'visual comfort zone'. At this time, several similar but not identical parameters have been identified, notably linked to the depth of field (±0.2 diopters), ±1 degree, +1/–2% [b-Chen], [b-Yano], [b-Speranza], [b-Mendiburu]. The visual comfort zone with these constraints depends on the viewing distance as depicted in the diagram of Figure 3.

The comparison of these different comfortable viewing zones is shown in Figure 3. Generally, these definitions generate similar comfort areas [b-Tam].

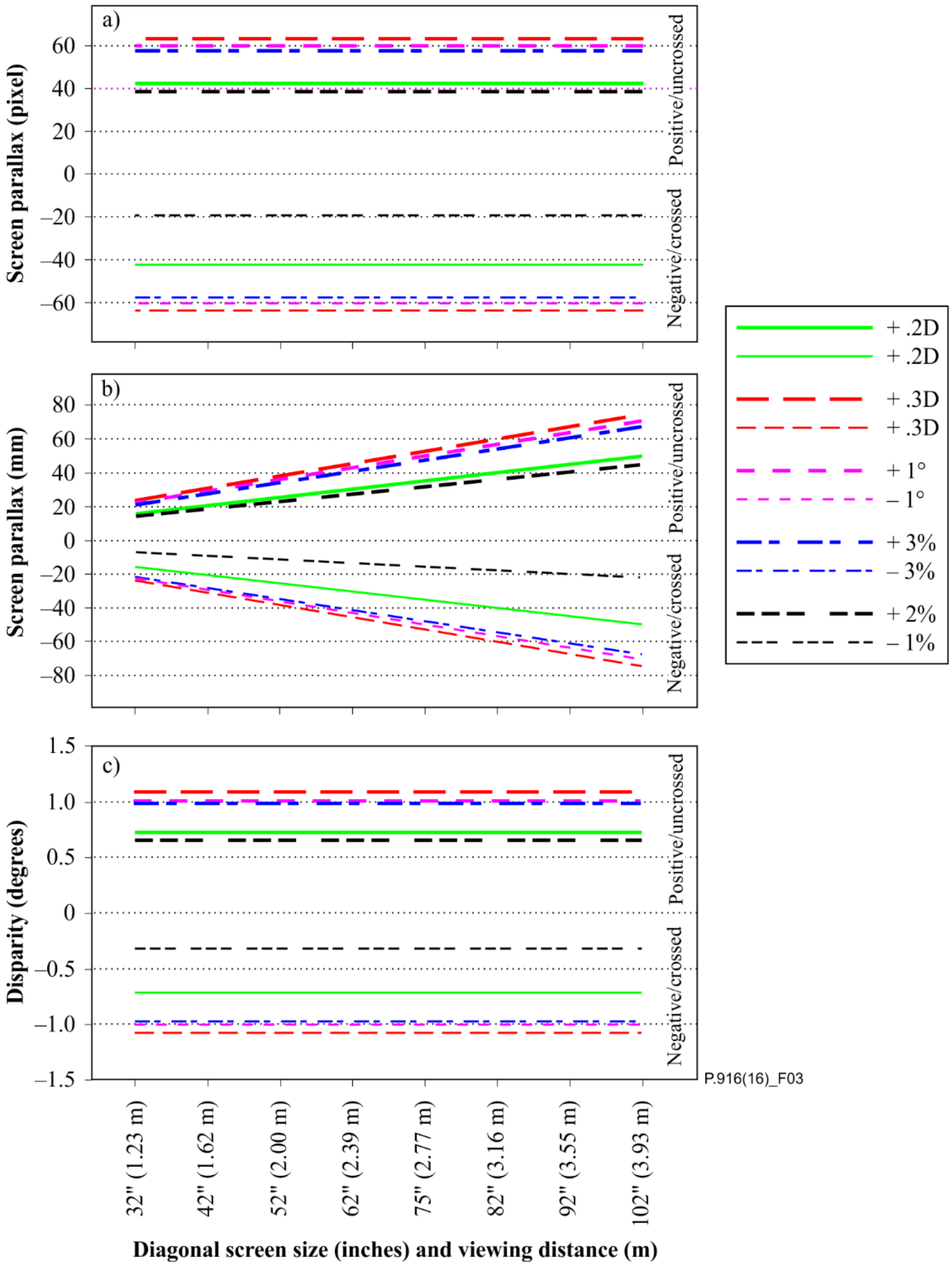


Figure 3 – Comparison on different definitions on comfortable viewing zone. The viewing distance is 3 times of the screen height [b-Tam]

7 Binocular disparity relationship in capturing, storing and rendering

This clause explains the link between the camera positions (capturing, the resulting disparity in the captured content and the perceived angular disparity at the viewer).

7.1 Conversion equations

The following equations are taken from [b-Chen].

Viewing distance

d: Viewing distance

p_x: pixel width

$$d = \frac{\frac{p_x}{2}}{\mathrm{atan}(\frac{\propto}{2})}$$

Maximum uncrossed disparities: (eyes should not diverge)

e: inter-pupillary distance

$$D_b^{max} = \frac{e}{p_x} D_b^{max} = \frac{e}{p_x}$$

Foreground distance of comfortable viewing zone

Z_f: largest distance between the farthest object in front of the display and the display

Z_b: largest distance between the farthest object in the background of the scene and the screen

$$Z_f = d - \frac{1}{\frac{1}{d} - 0.2} Z_b = \begin{cases} \frac{1}{\frac{1}{d} - 0.2} - d, & if\ d < 5 \\ ??\quad nn & otherwise \end{cases} Z_f = d - \frac{1}{\frac{1}{d} - 0.2}$$

$$Z_b = \begin{cases} \frac{1}{\frac{1}{d} - 0.2} - d, & if\ d < 5 \\ \infty, & otherwise \end{cases}$$

Depth rendering abilities in pixel

D_f: maximum crossed disparity in pixel

D_b: maximum uncrossed disparity in pixel

$$D_f = \frac{Z_f . e}{(d - Z_f) . p_w} D_b = \frac{Z_b . e}{(Z_b + d) . p_w} D_f = \frac{Z_f . e}{(d - Z_f) . p_w}$$

$$D_b = \frac{Z_b . e}{(Z_b + d) . p_w}$$

7.2 Conversion of pixel disparities to a perceptual space

7.2.1 Retinal disparities

Figure 4 represents the retinal binocular disparity, or stereopsis: the two eyes see two distinct objects and the projection of these objects on the retina appear to be at different locations on the retina of each eye. The differences of the position between these two projections are the retinal disparity. This information is processed by the brain to estimate of the relative position in depth between the two objects and is the most important binocular depth cue.

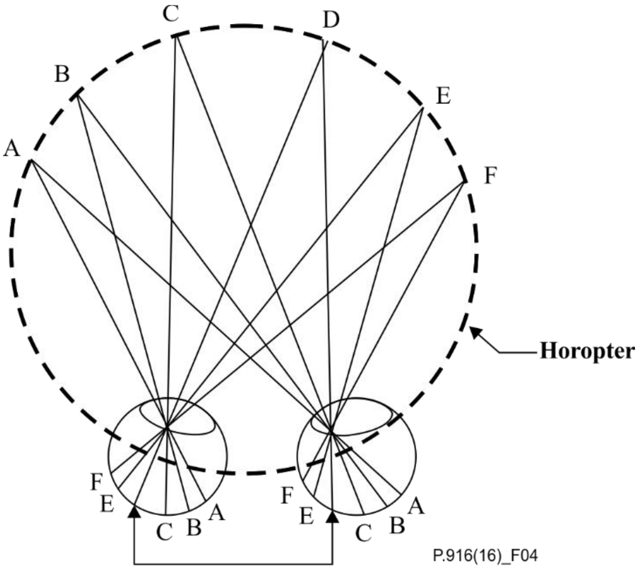


Figure 4 – Horopter

To determine the retinal disparities, from disparity in pixels on the display, the position of the 3D objects taking into account the viewing conditions can be determined using the following formula [b-Chen2]:

$$X = \frac{MBfx}{Bz + Mfb(1 - \frac{z}{d_{cov}})} Y = \frac{MBfy}{Bz + Mfb(1 - \frac{z}{d_{cov}})} Z = \frac{MBz}{Bz + Mfb(1 - \frac{z}{d_{cov}})} X$$

$$= \frac{MBfx}{Bz + Mfb(1 - \frac{z}{d_{cov}})}$$

$$Y = \frac{MBfy}{Bz + Mfb(1 - \frac{z}{d_{cov}})}$$

$$Z = \frac{MBz}{Bz + Mfb(1 - \frac{z}{d_{cov}})}$$

With:

f: the focal length of the camera

b: the inter-camera baseline (distance between focal length)

w_ccd,*h_ccd*: camera sensor width and height

d_cov: the convergence distance

(*x*,*y*,*z*): the position of a point in camera space (in front of the camera)

B inter-pupil baseline: the distance between the eyes of the observers

*W*_screen: the width of the display

M the magnification factor: M=W_screen/w_ccd

Based on the position of the object in the visualization space, it is possible to determine the retinal disparities using the set of equations proposed by Cormack [b-Cormack]. Figure 5 illustrates retinal disparities.

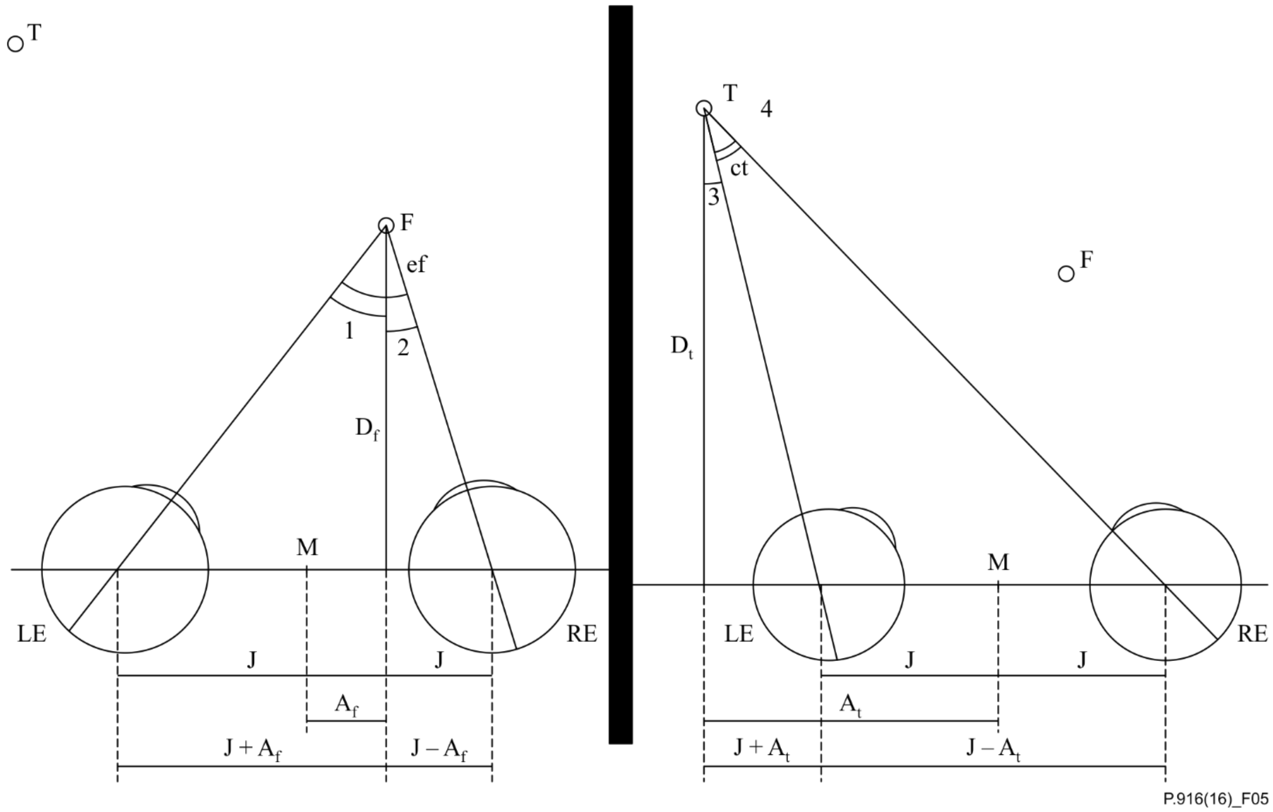


Figure 5 – Retinal disparities

The retinal disparities are the differences between two observed points:

$$r = angle_{cf} - angle_{ct} r = angle_{cf} - angle_{ct}$$

Each of these angles can be determined as follows:

$$\mathrm{angle}_{cf} \; = \; \mathrm{angle}_1 + \mathrm{angle}_2 \mathrm{angle}_{ct} \; = \; \mathrm{angle}_3 + \mathrm{angle}_4 \mathrm{angle}_{cf} \; = \; \mathrm{angle}_1 + \mathrm{angle}_2$$

$$angle_1 = \mathrm{atan}(\frac{\mathrm{e} + \mathrm{A_f}}{D_f})$$

$$angle_2 = \mathrm{atan}(\frac{\mathrm{e} - \mathrm{A_f}}{D_f})$$

$$\mathrm{angle}_{ct} \; = \; \mathrm{angle}_3 + \mathrm{angle}_4$$

$$angle_1 = \mathrm{atan}(\frac{\mathrm{e} + \mathrm{A_t}}{D_t})$$

$$angle_2 = \mathrm{atan}(\frac{\mathrm{e} - \mathrm{A_t}}{D_t})$$

Resulting in the following equation:

$$r = \left(\mathrm{atan}\left(\frac{\mathrm{e} + \mathrm{A_f}}{D_f}\right) + \mathrm{atan}\left(\frac{\mathrm{e} - \mathrm{A_f}}{D_f}\right)\right) - \left(\mathrm{atan}\left(\frac{\mathrm{e} + \mathrm{A_t}}{D_t}\right) + \mathrm{atan}\left(\frac{\mathrm{e} - \mathrm{A_t}}{D_t}\right)\right) r$$
$$= \left(\mathrm{atan}\left(\frac{\mathrm{e} + \mathrm{A_f}}{D_f}\right) + \mathrm{atan}\left(\frac{\mathrm{e} - \mathrm{A_f}}{D_f}\right)\right) - \left(\mathrm{atan}\left(\frac{\mathrm{e} + \mathrm{A_t}}{D_t}\right) + \mathrm{atan}\left(\frac{\mathrm{e} - \mathrm{A_t}}{D_t}\right)\right)$$

In the special case of the midsagittal plane, it is possible to simplify these equations:

$$r = 2 \times \mathrm{atan}\left(\frac{e}{Df}\right) - 2 \times \mathrm{atan}(\frac{e}{D_t})$$

7.2.2 Parallax in degree

Parallax values can be computed to study the perceptual properties of the source content under study enabling selection of 3D content while being aware of their depth characteristics.

Parallax values can be determined as follows [b-Lin]:

$$P = a - ba = \tan^{-1}\left(\frac{D_{IP} + D_s - 2T_s}{2L}\right) + \tan^{-1}\left(\frac{D_{IP} + D_s + 2T_s}{2L}\right) b$$
$$= \tan^{-1}\left(\frac{D_{IP} + D_s}{2L}\right) + \tan^{-1}\left(\frac{D_{IP} + D_s}{2L}\right) P = a - b$$

$$a = \tan^{-1}\left(\frac{D_{IP} + D_s - 2T_s}{2L}\right) + \tan^{-1}\left(\frac{D_{IP} + D_s + 2T_s}{2L}\right)$$
$$b = \tan^{-1}\left(\frac{D_{IP} + D_s}{2L}\right) + \tan^{-1}\left(\frac{D_{IP} + D_s}{2L}\right)$$

With (Figure 6):

D_IP: the inter-pupillary distance in metres

D_s: the disparity on the display in metres

L: the viewing distance

T_s: determine the convergence point which considering the surface of the horopter may not be constant in the plane screen [b-Schreiberetal]

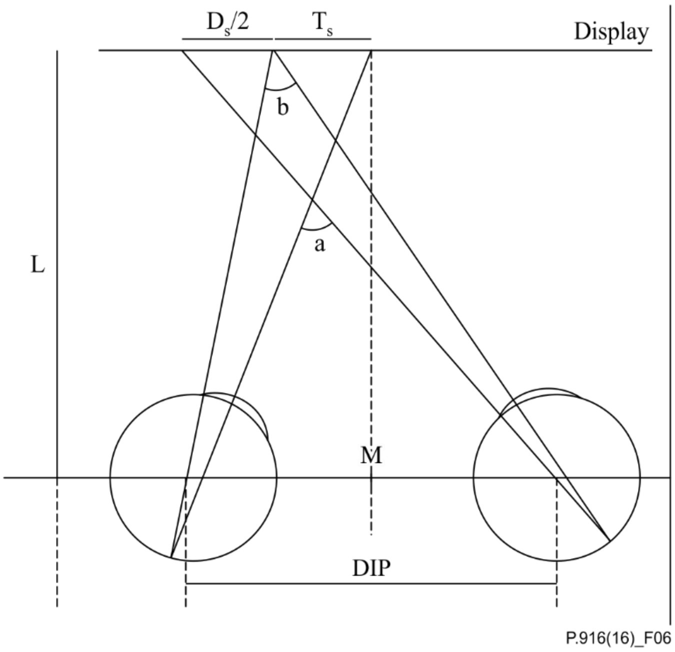


Figure 6 – Parallax value computation.

7.3 Typical constraints for displays

As expressed though the equations in clause 7.1, the use of different displays having different resolutions and sizes, in addition to different viewing conditions affects the 3D rendering capabilities. Table 1 applies the equations from clause 7.1 to typical display configurations, namely standard definition (SD), high definition (HD) and head-up display (HUD), providing information for these standard scenarios of what the viewing distance should be to ensure the perception of the full resolution of the display. The viewing distance being set, this implies:

– Constrains on the contents maximum disparities to ensure a comfortable viewing experience

– A depth range where the 3D object will be perceived

Table 1 – Viewing distances and comfortable viewing zones

	SD			
Resolution	720x576		720x480	
Display diagonal (inch)	24	32	24	32
Display height (m)	0.366	0.488	0.366	0.488
Pixel width (mm)	0.635	0.847	0.762	1.016
Viewing distance (m)	2.183	2.911	2.620	3.493
Viewing distance (times H)	5.968	5.968	7.162	7.162
Maximum uncross disparity (pixel)	102	77	85	64
Foreground dist. comfortable viewing zone (m)	0.663	1.071	0.901	1.436
Background dist. comfortable viewing zone (m)	1.692	4.055	2.883	8.094
Number of pixel in the comfortable zone (foreground)	44.7	44.7	44.7	44.7
Number of pixel in the comfortable zone (Background)	44.7	44.7	44.7	44.7

Table 1 – Viewing distances and comfortable viewing zones (cont.)

	HD					
Resolution	1920x1080			1920x540		
Display diagonal (inch)	40	46	55	40	46	55
Display height (m)	0.498	0.573	0.685	0.275	0.316	0.378
Pixel width (mm)	0.461	0.530	0.634	0.509	0.586	0.700
Viewing distance (m)	1.586	1.823	2.180	1.751	2.014	2.408
Viewing distance (times H)	3.183	3.183	3.183	6.366	6.366	6.366
Maximum uncross disparity (pixel)	141	123	102	128	111	93
Foreground dist. comfortable viewing zone (m)	0.382	0.487	0.662	0.454	0.578	0.783
Background dist. comfortable viewing zone (m)	0.736	1.047	1.685	0.944	1.358	2.237
Number of pixel in the comfortable zone (foreground)	44.7	44.7	44.7	44.7	44.7	44.7
Number of pixel in the comfortable zone (Background)	44.7	44.7	44.7	44.7	44.7	44.7

Table 1 – Viewing distances and comfortable viewing zones (end)

	UHD		
Resolution	3840x2160		
Display diagonal (inch)	40	46	55
Display height (m)	0.498	0.573	0.685
Pixel width (mm)	0.231	0.265	0.317
Viewing distance (m)	0.793	0.912	1.090
Viewing distance (times H)	1.592	1.592	1.592
Maximum uncross disparity (pixel)	282	245	205
Foreground dist. comfortable viewing zone (m)	0.108	0.141	0.195
Background dist. comfortable viewing zone (m)	0.149	0.203	0.304
Number of pixel in the comfortable zone (foreground)	44.7	44.7	44.7
Number of pixel in the comfortable zone (Background)	44.7	44.7	44.7

8 3D video content guidelines

8.1 Introduction

The goal of 3D video is to enhance the perception of reality by altering depth cues, in particular by introducing binocular depth cues.

The overall sensation of reality a viewer sees depends on technical factors that impact perceptual dimensions. The technical factors include resolution and colour-fidelity. The perceptual dimensions include sharpness and depth cues.

This affects the potential for the sensation of reality, but the actual sensation of achieved reality depends also on the combination of the technical factors and the scene content itself. For example, if all of the objects in a wide angle shot are far away from the viewpoint, then the scene will generally not benefit from binocular depth cues, because those are not used by the viewer in this situation.

8.2 Disparity constraints

Excessive disparity/parallax possibly causes visual discomfort, because it worsens the conflict between accommodation and vergence. Therefore, it has been suggested that to minimize the accommodation-vergence conflict, the disparities in the stereoscopic image should be small enough so that the perceived depths of objects fall within a "comfort zone".

The stereoscopic test content should normally be comfortable to watch. The visual comfort of stereoscopic images depends critically upon the disparity contained in the image. The visual comfort of stereoscopic images also depends upon the subject's inter-pupillary distance and the viewing distance.

Figure 7 shows the relationship between inter-pupillary distance, viewing distance, content disparity and depth range.

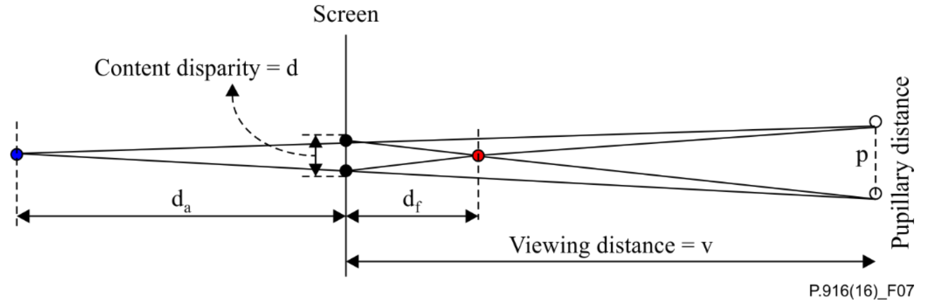


Figure 7 – Relationship between inter-pupillary distance *p*, viewing distance *v*, content disparity *d* and depth range (d_f, d_a)

Inter-pupillary distance is usually set as 6.5 cm.

Content disparity can be positive or negative. The original content disparity is described by the unit of *pixels*. However, due to the differences on the displaying screen size, the real distance of the disparity will be changed when using the unit of centimetres. Positive values mean the eyes will be converged behind the screen (uncrossed disparity). Negative values mean the eyes will be converged in front of the screen (crossed disparity).

Depth range is the distance of the converged point (virtual object) in front of the screen and behind the screen.

$$d_f = \frac{d \cdot v}{p + d}$$

$$d_a = \frac{d \cdot v}{p - d}$$

Even though displaying the same 3D content, as described above, the screen size will change the depth range that observers perceive. Figure 8 shows three different screens with sizes of 24", 40" and 46". The viewing distance is 3H for all cases. The pupillary distance is set as 6.5 cm. The 3D content disparity is ±30 pixels.

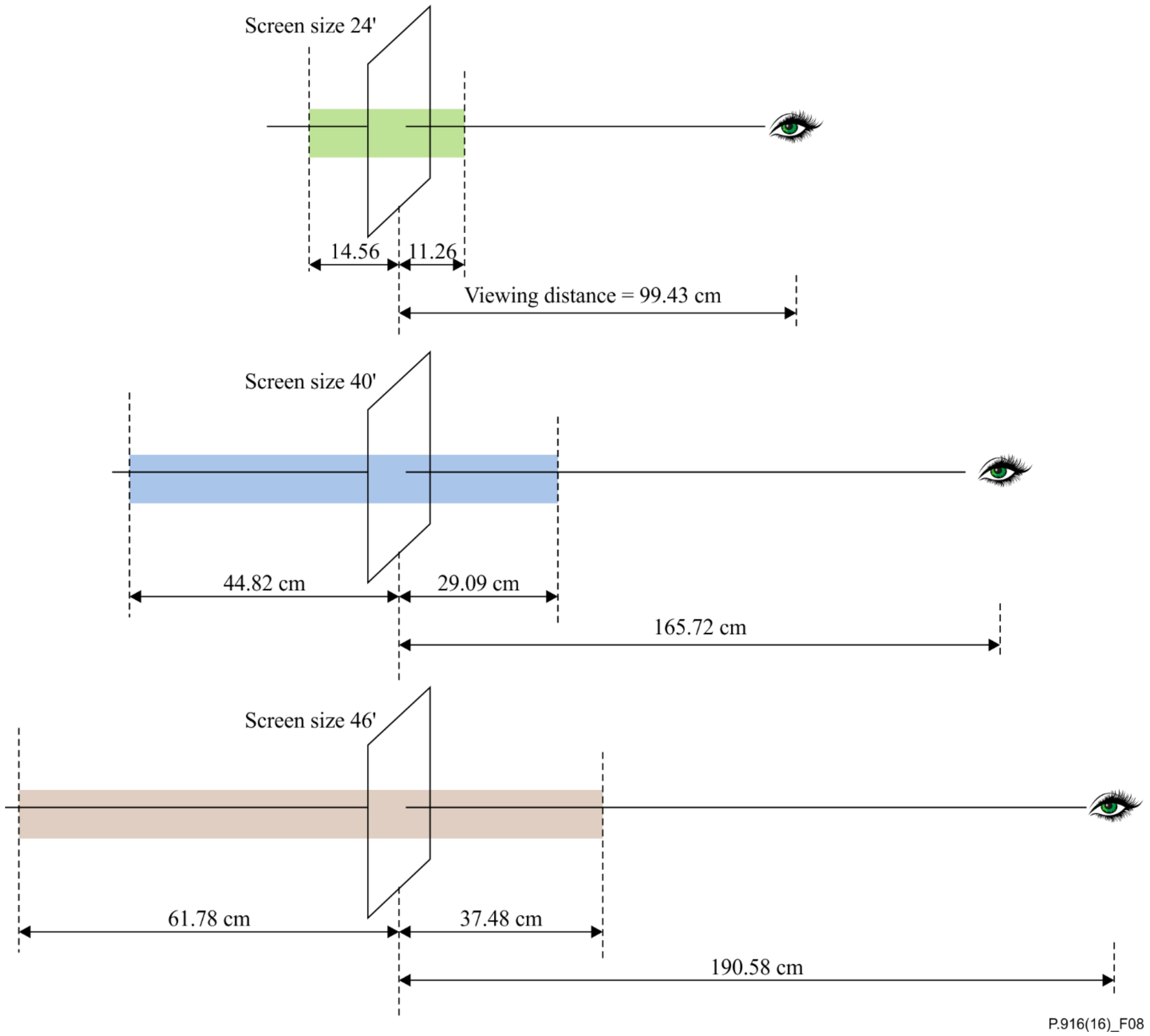


Figure 8 – Depth ranges as a function of screen sizes

The following safety guidelines are recommended for filming, production and selection of stereoscopic 3D video content.

- Fast spatial and temporal changes of disparity in stereoscopic 3D content should be avoided. The fast change of binocular disparity may induce visual discomfort accompanied by symptoms such as focusing difficulty and eye strain.
- To prevent visual discomfort, excessive disparity of stereoscopic 3D content should be avoided. Excessive binocular disparity may induce visual discomfort accompanied by symptoms such as focusing difficulty and eye strain.
- Image content with rotating or lateral shaking may induce the sensation of self-motion (visually induced self-motion sensation).

8.3 View asymmetries

Viewers may not feel comfortable when viewing left and right images that have size, verticality, inclination and brightness differences. Cross-talk between the left and right images may also have an impact on viewing comfort.

These safety guidelines may be ignored for brief periods of time, for creative purposes.

Thresholds for disparity should be expressed in angular degrees of the viewing angle.

All 3DTV systems displayed on a screen in a single plane (such as a television screen) have limitations for a number of reasons. One of them is the potential conflict between convergence (the object that the eyes point themselves towards) and accommodation (the point on which the lens of the eye focuses) which the two signals gives rise to. 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 3D 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 3D systems. Optimizing 3D systems is the process of minimizing the effects of the limitations.

Visual fatigue caused by viewing stereoscopic motion images is a particular safety concern. Viewers' repeated adaptation to the discrepancy between eye convergence and accommodation causes a decline of their visual functions and results in visual fatigue.

8.4 Scene cuts

Attention should be paid to the distribution of parallaxes in the stereoscopic images. From the correlations between psychological factors and the parallax distribution, we can grasp the essential characteristics of stereoscopic images, e.g., the sense of presence they convey and their ease of viewing. The parallax distribution of stereoscopic images is discontinuous during scene-change frames, where the scene depth and perceived convergence distance change. We need to evaluate how these changes affect the visual discomfort experienced during viewing of stereoscopic images.

8.5 Excessive disparity/parallax

Excessive disparity/parallax would cause visual discomfort. A possible reason is that if the disparity is large enough, the conflict between accommodation and vergence will be increasing. In addition, the observer may feel annoyance due to the difficulties in fusing the left and right eye images. This occurs as a consequence of the fusion limit, which will introduce "double vision".

Therefore, it has been suggested that, to minimize the accommodation-vergence conflict and to avoid "double vision", the disparities in the stereoscopic images should be small enough so that the perceived depths of objects can fall within the "comfortable viewing zone" defined in clause 6.3.

Excessive 3D pop-out effects in content shall be limited in the number of times of occurrences and duration.

9 3D viewer guidelines

9.1 The influence of stereopsis and abnormal binocular vision (ABV)

Subjects with abnormal binocular vision (ABV) such as strabismus, amblyopia and an isometropia have decreased 3D perception (e.g., more difficulty perceiving 3D).

Among people with ABV, subjects that have a normal degree of stereopsis are more susceptible to 3D fatigue than subjects that have poor stereopsis. This includes 3D fatigue symptoms such as dizziness, headaches, eye fatigue and pain. A normal degree of stereopsis is for example good stereopsis as measured with a Titmus stereofly test.

If a person cannot perceive 3D, or feels severe 3D fatigue while watching 3D content, he or she should consult an ophthalmologic specialist for evaluation of abnormal binocular vision. However, badly produced 3D content will cause 3D fatigue in everyone.

Visual functions vary greatly from person to person, so it is essential to understand that there are individual differences before subjective assessment begins. For instance, there are limits to the binocular parallax of left and right images which a person can fuse into one image; when the parallax exceeds these limits, a double image is perceived. In this situation, depth perception collapses and viewing becomes extremely uncomfortable. For this reason, it is necessary to know the range of binocular parallax over which two images can be fused into one. However, individual differences are vast and will necessitate a study of the stereopsis function of many people.

9.2 Warnings for subjects

Viewers of 3DTV should be made aware of the following warnings:

- Avoid viewing 3D TV if body functions start to deteriorate or if excessive stress is perceived.
- Stop viewing 3D TV if double vision is experienced.
- Image content with rotating or lateral shaking may induce the sensation of self-motion (visually induced self-motion sensation). Stop watching 3D TV temporarily if you feel uncomfortable with this sensation.
- Individuals with a shorter inter-pupillary distance may feel more visual discomfort with 3D TV than others (e.g., children).
- Individuals who experience motion sickness in automobiles or amusement park rides may experience more visual discomfort during 3D TV viewing.

9.3 Symptoms of 3D visual fatigue

The following symptoms may indicate 3D visual fatigue. More information can be found in [b-Kuze].

- Difficulty in tracking the motion on TV.
- Eye strain: blurred vision, bleary, dry eyed, gritty, eye ache, sting, eyes heavy, hazy, warm eyes, flickering and watery eyes.
- Focusing difficulty: double vision, near vision difficulty, far vision difficulty and trouble fusing stereoscopic images.
- General fatigue symptoms: feeling heavy in the head, difficulty in concentration, dizziness, stiff shoulder and stiff neck.
- Headache: pain in the temple and pain in the middle of the forehead.
- Nausea: vomiting and vertigo.
- Transient visual dimness after watching TV

If a person feels any of these symptoms, it may be an indication of visual fatigue. He or she should stop 3DTV viewing until the side effect disappears.

If a person repeatedly experiences 3D fatigue symptoms, he or she should consult ophthalmologic specialists for evaluation of abnormal binocular vision. The 3D TV hardware operation should also be examined for problems.

10 3D viewing environment and playback guidelines

Illumination, sound, viewing height and TV brightness and focus should be adjusted to comfortable levels when watching 3D TV.

If an image appears as a double image or if it is hard to perceive a 3D image, stop using the system and check for errors in the display settings. If you still experience stereopsis problems with no apparent hardware problems, then you should temporarily stop viewing 3D TV.

Some viewers might not notice whether left and right images are reversed. The reversed vision of left and right images can cause visual fatigue and discomfort. Reversed left and right images can be caused by many problems (e.g., editing errors, hardware problems, camera setup, viewing angle).

10.1 Relationship between viewer and 3D monitor

The viewer should be seated as advised by the 3D television manual, at a comfortable distance from the monitor. For television monitors, a minimum of 3H (three times picture height) is advised. For computer monitors, a minimum of 1.5H is advised.

Optimal 3D viewing means the viewer should be seated upright, in front of the screen (i.e., minimal head tilt). The viewing direction should be normal (e.g., head facing perpendicular) to the screen. The horizontal viewing angle from the screen should be minimum (e.g., face the screen directly instead of at an angle).

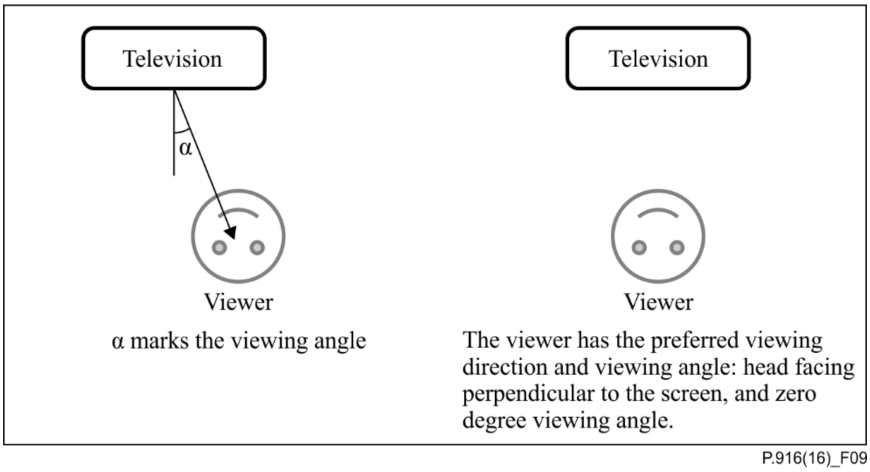


Figure 9 – Viewing direction and viewing angle

Figure 9 demonstrates viewing direction and viewing angle. It is generally said that the minimum value for depth of field of the human eye is ±0.3 D, where diopter (D) is the reciprocal value of distance (m). This means that we can perceive the image without defocusing when the object is located within ±0.3 D. When viewing 3D television, the accommodation point is fixed on the screen and therefore 3D pictures should preferably be displayed within this range. Since ordinary television programs include images at infinite distance, the desirable range of depth to be displayed with 3D 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 3D 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.

The effective viewing angle should allow 20% angular rotation of head movement in the horizontal plane.

10.2 Viewing time and visual fatigue

Acceptable viewing time for 3D content varies depending upon the content. A short time resting is recommended for every 1 or 2 hours, or when the eyes feel tired.

11 Assessment methods for symptoms of visual fatigue

11.1 Questionnaire to assess symptoms of visual fatigue

Appendix II provides a questionnaire to assess the symptoms of visual fatigue. This questionnaire is taken from [b-Kennedy].

For an example of this technique, see [b-Brunnström].

11.2 Pair comparison (PC) subjective test to visual discomfort

In [ITU-T P.913], the pair comparison method has been well defined. In pair comparison, the test sequences are presented in pairs, consisting of the same sequence being presented first through one system under test and then through another system. In 3DTV conditions, generally, we compare two sequences with the same content but different distortions, or disparity setting, etc.

The pair comparison method requires that for the test sequences, for example, N different hypothetical reference circuits (HRCs), the observers have to compare all the possible pairs, which leads to N(N–1) pairs for one content (if considering the influence of presentation order) or N(N–1)/2 pairs (if without considering the presentation order).

To avoid the influence of presentation order, all the pairs of sequences must be displayed in both the possible orders (e.g., AB, BA). This can be implemented by one observer (the observer watches AB and BA) or by a group of observers, which means, for one observer, if he watches pair AB, there must be another observer who watches the pair with the order BA.

After each pair a judgement is made on which element in a pair is preferred in the context of the test scenario. For the assessment of visual discomfort or visual fatigue, the question to observer may be "which one is more comfortable/uncomfortable?".

In paired comparison tests, there are two presentation patterns for the stimulus pair, i.e., the way to display the stimulus pair to observers. The stimuli can either be presented one after the other on a single screen ("Time-sequential presentation") or they can be presented on two well calibrated and synchronized screens ("Time-parallel presentation"). Different presentation patterns can be used for different experimental setups.

The time pattern for the stimulus presentation can be illustrated as in Figure 10 and Figure 11. The voting time should be less than or equal to 10 s, depending upon the voting mechanism used. The presentation time should be about 10 s and it may be reduced or increased according to the content of the test material.

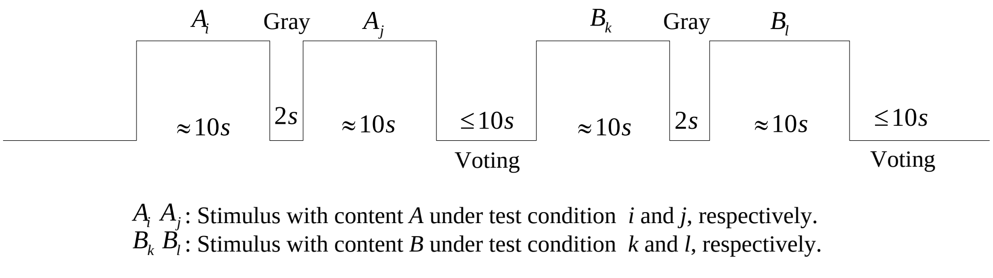
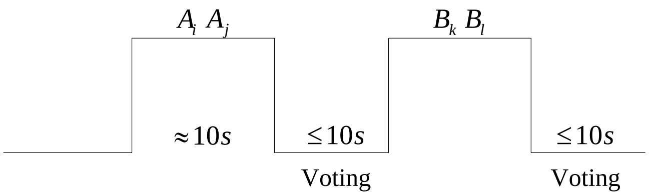


Figure 10 – Time pattern for time-sequential pair comparison presentation



A_i A_j: Stimulus with content *A* under test condition *i* and *j*, respectively.
B_k B_l: Stimulus with content *B* under test condition *k* and *l*, respectively.

Figure 11 – Time pattern for time-parallel pair comparison presentation

Appendix I

Open questions

(This appendix does not form an integral part of this Recommendation.)

More research is needed on the accommodation-vergence conflict. [b-Shiomi] indicates that there is no accommodation-vergence conflict.

More research is needed on the relationship between viewing distance and monitor type (e.g., active vs passive).

Appendix II

Visual fatigue questionnaire

(This appendix does not form an integral part of this Recommendation.)

This appendix provides an example questionnaire. The questions may be adapted as to the requirements of the specific experiment.

The following questionnaire on visual fatigue is conducted in two phases. The first set of questions is answered by observers prior to the subjective assessment in order to draw their attention to certain effects:

1) Are you blinking your eyes more frequently than usual today? (No, slightly, a bit, to a larger extent, much more frequently than usual)
2) Do you have difficulties with your near vision today (Reading, cooking, sewing)? (No, slightly, a bit, to a larger extent, much more)
3) Do you have difficulties with night vision? (No, slightly, a bit, to a larger extent, a lot)
4) Do you have double vision? (No, slightly, a bit, to a larger extent, a lot)
5) Do you feel pain at the back of your head? (No, slightly, a bit, to a larger extent, a lot)
6) Do you have problems focusing? (No, slightly, a bit, to a larger extent, a lot)
7) Do you have difficulties driving at night? (No, slightly, a bit, to a larger extent, a lot)
8) Are you worried about your vision? (Never, Rarely, Sometimes, Often, Always)
9) Are your eyes watering? (No, slightly, a bit, to a larger extent, a lot)
10) At this moment, your two-eyed vision is: (Excellent, very good, good, fair, poor)
12) Do you have dry eyes? (No, slightly, a bit, to a larger extent, a lot)
12) Do you have a stiff neck? (No, slightly, a bit, to a larger extent, a lot)

After the subjective experiment, the following questions are answered:

1) Do you have the sensation that the movement of your eyes is decoupled? (yes, no)
2) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) stiffness of your neck?
3) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) eye strain?
4) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) dizziness?
5) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) lightheaded?
6) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) heavy eye lids?
7) Did your eyes water during the experiment? (yes/no)
8) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) stabbing pain in your eyes?
9) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) pain at the front of your head?

10) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) fatigue in the eyes?

11) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) problems concentrating?

12) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) sleepiness?

13) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) pain at your temples?

14) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) nausea?

15) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) blurred vision?

16) Do you have the feeling that your eyes are looking in different directions (yes/no)

17) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) eye blinking?

18) During the experiment, did you close your eyes in order to re-establish a clear vision (yes/no)

19) Do you experience double vision (yes/no)?

20) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) pain at the back of your head?

21) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) difficulties focusing?

22) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) watering of the eyes?

23) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) dry eyes?

24) With respect to the beginning of the experiment, do you experience (a lot more, slightly more, a little bit more, the same, a little bit less, slightly less, a lot less) pain in your shoulders?

25) During the experiment, did you need to look at a different object than the screen? Which one? (Yes – object, No)

Bibliography

[b-Brunnström] Brunnström, K., Wang, K., and Andrén, B. (2013), *Simulator sickness analysis of 3D video viewing on passive 3DTV*, Stereoscopic Displays and Applications XXIV, Vol: 8648, A. J. Woods, N. S. Holliman and G. E. Favolora Eds, paper 1.

[b-Chen] Chen, W., Fournier, J., Barkowsky, M. & Callet, P.L. (2010), *New Requirements of Subjective video quality assessment methodologies for 3DTV, Video Processing and Quality Metrics 2010 (VPQM)*, Scottsdale, USA.

[b-Chen2] Chen, W., Fournier, J., Barkowsky, M. & Callet, P.L. (2011), *New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone*, Stereoscopic Displays and Applications XXII. Proceedings of the SPIE 7863, 78631O-78631O-13.

[b-Cormack] Robert Cormack (1985), *The computation of retinal disparity*, *Perception and Psychophysics*, vol. 37, No. 2, pp. 176-178.

[b-2293] Report ITU-R BT.2293-1, *Principles for the comfortable viewing of stereoscopic three-dimensional television (3DTV) images*.

[b-Kennedy] Kennedy, R.S, Lane, N.E, Berbaum, K.S and Lilienthal, M.G (1993), *Simulator Sickness Questionnaire: An Enhanced Method of Quantifying Simulator Sickness*, The International Journal of Aviation Psychology 3 (3), 203-220.

[b-Kuze] Kuze, J, & Ukai, K. (2008). *Subjective evaluation of visual fatigue caused by motion images*, *Displays*, 29(2), 159-166.

[b-Lin] T.-H. Lin and S.-J. Hu (2014), *Perceived depth analysis for view navigation of stereoscopic three-dimensional models*, Journal of Electronic Imaging, vol. 23, No. 4, p. 043014.

[b-Mendiburu] Mendiburu B. (2009), *3D movie making : stereoscopic digital cinema from script to screen*, Focal Press, Burlington, MA, Etats-Unis.

[b-Schreiberetal] Schreiberetal, K.M, *Thesurfaceoftheempiricalhoropter* (2008), Journal of Vision 8(3), 1-20.

[b-Shiomi] Shiomi, T, Hori, H, Uemoto, K, Miyao, M, Takada, H, Hasegawa, S., Omori, M, Watanabe, T and Ishio, H (2012), *Comparison of Simultaneous Measurement of Lens Accommodation and Convergence in Natural Vision and 3D Vision*, *SID Symposium Digest of Technical Papers*, Vol: *43*, 1178-1181.

[b-Speranza] Speranza F., Tam W.J, Renaud R., Hur N. (2006), *Effect of disparity and motion on visual comfort of stereoscopic images*, Proceedings of SPIE Stereoscopic Displays and Virtual Reality Systems XIII, vol. 6055, San Jose, CA, Etats-Unis, p. 94-103.

[b-Tam] W.J. Tam, F. Speranza, S. Yano, K. Shimono and H. Ono (2011), *Stereoscopic 3D-TV: visual comfort*. Broadcasting, IEEE Transactions on, 57(2):335-346.

[b-Woods] Andrew Woods, Tom Docherty and Rolf Koch (1993), *Image distortions in stereoscopic video systems*, *Proceedings of the SPIE, Stereoscopic Displays and Applications IV*, vol. 1915, pp. 36-48.

[b-Yano] Yano S., Ide S., Mitsuhashi T., Thwaites H. (2002), *A study of visual fatigue and visual comfort for 3D HDTV/HDTV images*, Displays, vol. 23, No. 4, p. 191-201.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	**Terminals and subjective and objective assessment methods**
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2016