International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Report ITU-R BT.2160-3
(06/2012)

Features of three-dimensional television video systems for broadcasting

BT Series

Broadcasting service (television)

International Telecommunication Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Reports

(Also available online at http://www.itu.int/publ/R-REP/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	**Broadcasting service (television)**
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management

***Note**: This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2012

REPORT ITU-R BT.2160-3

Features of three-dimensional television video systems for broadcasting

(2009-2010-2011-2012)

Page

Summary

The technology needed for a first-generation three-dimensional television (3DTV) two-channel stereoscopic system already exists, although so far there have been no announced plans for the general introduction of regular free-to-air broadcasting services. A number of broadcasting organizations nevertheless continue to carry out experiments in stereoscopic 3DTV production, while pay-television operator BSkyB introduced a stereoscopic 3DTV channel in the United Kingdom in October 2010. Several consumer electronics manufacturers introduced stereoscopic television receivers during 2010.

An essential aim of this Report is to present a framework for a study of the various aspects of digital three-dimensional (3D) TV broadcasting systems[1] as outlined in Question ITU-R 128/6. It is intended to identify the issues that need to be addressed, and to encourage further contributions to WP 6C.

1 Motivations for the introduction of 3DTV broadcasting

Interest in the possibility of 3DTV in the home may be due in part to a new wave of 3D movies reaching the cinema. In spite of the need to wear glasses, 3D movies have proved to be popular, attracting large audiences who are prepared to pay a premium for the 3D experience.

This in turn has created expectations of the imminent arrival of 3D movies in the home through packaged media[2], such as DVD and Blu-ray. Movies are an important part of television broadcasting, and so it is natural to consider whether 3D movies might in due course be made available through broadcast means.

On the other hand, while the need to wear glasses has not been an impediment to the success of 3D-cinema, questions are raised about the suitability of glasses in the home environment. The current state of development of autostereoscopic displays for glasses-free viewing leaves much to be desired, although it is hoped that ongoing research will eventually lead to improved or even new forms of glasses-free display.

So today's motivation to explore the possibility of the introduction of 3DTV broadcasting may be seen partly as exploitation of the natural evolution of the phased delivery chain used for movies where feature films are first screened in the theatre, then go to the home in packaged media, and finally are made available on broadcast television. In addition, a pay television operator may also have an interest in offering premium content in 3D, whether movies or live events.

Lastly, although 3DTV might not currently be seen a "future alternative" or development of high–definition television (HDTV), it is certainly possible that it could at least have a complementary role to other forms of 3D experience that are likely to become available in the home in the not too distant future.

---

[1] Digital (3DTV) broadcasting is a television system that is designed to convey by broadcast transmission a more natural impression of depth to the scene that is being portrayed, by rendering spatially different views to each eye.

In its simplest form, the viewer is presented with a fixed or "static" stereoscopic view of the scene, while a more complex form of 3DTV enables the viewer to change the perspective of what is seen in discrete steps, by means of head movement. In its ultimate form, "holographic" three dimensional presentation would enable the viewer to change perspective through head movement in a continuous way that is comparable to natural sight in real life.

[2] A Blu-ray 3D specification has been introduced that encodes 3D video using the 'Stereo High' profile defined by Multiview Video Coding (MVC), an extension to the ITU-T H.264 Advanced Video Coding (AVC). This is able to provide full 1080p resolution per eye, and can offer backward compatibility with existing 2D Blu-ray Disc players.

2 Background to possible 3DTV systems

The fundamental means by which a 3DTV broadcast system today is capable of enhancing the user's visual experience of three-dimensionality, compared to the broadcast of HDTV images, is by delivering stereoscopic image information to viewers in the home. 3DTV broadcasts must provide the signals necessary for generating images with different views of a scene to the two eyes of a viewer. By means of binocular fusion of the stereoscopic images, the 3DTV viewer can obtain an enhanced sensation of depth and an improved sensation of "presence" and "reality".

It is envisioned that the technology of 3DTV systems, as with all media systems, will develop and advance from one generation to the next, over a period of possibly many years. It may be anticipated that future generations will be likely to increase the amount of visual information provided, reduce the restrictive need for eyewear, and increase the freedom of movement allowed without negatively affecting the quality of the stereoscopic depth.

Thus, one method of classifying the various 3DTV systems is as follows:

Eyewear-based systems:

Those systems that are based on or targeted for "plano-stereoscopic" displays, whereby left and right eye images are presented independently to the two eyes using various methods that require eyewear to isolate the two views of a given scene.

Multiview autostereoscopic systems:

Such systems that are targeted for "plano-stereoscopic" (or non-volumetric) displays whereby left and right eye images are presented independently to the two eyes, using various methods that allow two views of a given scene to be isolated without the need for eyewear. In addition, this generation of systems may provide multiple views of a scene such that viewers can freely change their viewing angle and have access to visual scene behind objects.

Integral imaging or holographic system:

Those systems that are based on object-wave recording (holography) or integral imaging and are targeted at the simulation of a light field generated by an actual scene. Thus, freedom of viewing position without the hindrance of eyewear is provided. In addition, the light field provides the visual information (focus cues) for adjusting the ocular lens so as to focus correctly at the same distance as the convergence distance. This provides more natural viewing than the systems of the previous generations that requires maintenance of focus at the display screen irrespective of convergence distance.

3 A hierarchical structure

Current proposals for 3DTV signal formats can be seen as forming a hierarchical structure, which correspond to different constraints and requirements. This is given in diagrammatic form in Fig. 1. This hierarchy might be used in future for any draft Recommendation for 3DTV by the ITU-R found to be required.

FIGURE 1

Matrix of signal formats for 3DTV

Compatibility level

Conventional HD Service Compatible (CSC) **Level 4**	**2D HD + MVC [(1)] (i.e. MVC)** (L, R formed by matrixing)	**2D HD + MVC [(2)] (i.e. MVC)** (Depth, occlusion, transparency data)	
HD Frame-Compatible Compatible (FCC) **Level 3**	Frame compatible plus MPEG resolution extension, for example SVC [(3)]		
Conventional HD Frame Compatible (CFC) **Level 2**	Frame compatible (L, R in same HD frame)		
Conventional HD Display Compatible (CDC) **Level 1**	Optimized colour anaglyph		
	Plano-stereoscopic profile **1st generation 3DTV**	Multiview profile **2nd generation 3DTV**	Object wave profile **3rd generation 3DTV**

Generation profile

Report BT.2160-01

(1) MPEG-4 AVC Stereo High Profile – a subset of MPEG Multiview Video Coding (MVC), Annex H to Recommendation ITU-T H.264.
(2) It is understood that ISO/IEC JTC/SC29/WG11 intends to address this form of extension to MVC.
(3) Annex G to Recommendation ITU-T H.264. Use of SVC would require a Level 5 decoder; Level 5 does not support interlace.

The principle of the hierarchy is that each box in the matrix in the diagram defines a type of signal, and this would correspond to the needs of a generic type of receiver. This is somewhat similar to the concept used for ISO/IEC JTC1 MPEG standards, though there are differences. Upper levels are intended to be “backward compatible” with lower levels, with one exception which is explained later.

Though different 3DTV display technologies today have different advantages and disadvantages, the hierarchy is essentially independent of the type of display used. Research and development, and market forces should allow 3DTV displays and technology to evolve and improve, while preserving the public interest for interoperability.

The hierarchy needs to cope with a range of circumstances, from where existing receiving equipment must be used intact (though glasses are used), to where some new elements (displays) are acceptable, to where both new receivers and displays are acceptable.

The quality of the 3DTV will be influenced by the quality of the individual left-eye and right-eye signals, and because of this, 3DTV may be most effective for the higher quality environment rather than the SD-TV environment.

Broadcasters may choose to use available 3DTV technology, and find the limitations acceptable, bearing in mind the gains, or they may prefer to wait for future technology which will have fewer limitations. It seems desirable that ITU-R should provide guidance for both.

3.1 Technology generations

In Fig. 1, the x-axis relates to the **system "generation"**. We may expect basic 3DTV technology to evolve in the decades ahead. The pattern of evolution is that we move from viewing a single stereo view with glasses, then to viewing with greater freedom for head movement, finally to viewing as we do normally ("natural vision").

Broadcasters may decide to begin broadcasting with earlier technology generations (with its limitations), or to wait for future generations.

First-generation technology is based on the capture and delivery of two views, one for the left eye, and one for the right eye. There is a single "binocular disparity" or binocular parallax. There are limitations with such systems, compared to "natural vision". With careful production, delivery, and display, effective results can be achieved. Usually, special glasses are used for viewing, though viewing without glasses (auto-stereoscopic) viewing is also possible.

Second generation technology is based on capture and delivery of multiple views. This allows multiple binocular disparities which makes the viewing experience closer to "natural vision". Normally viewing will be done without glasses on auto-stereoscopic displays.

Third generation technology is based on the capture and delivery of the "object wave", as is done in a simple way today with holography. The development of such systems is many years away at the moment.

We cannot predict with certainty whether, when, or if, the higher generations will be developed. But, we may note that often generation steps occur about every ten years or so, and that there can be a long lead time from idea to commercial exploitation.

3.2 Compatibility levels

The levels, or y-axis, in Fig. 1 relates to **compatibility levels**.

Level 1 relates to signals which provide for a system which does not require any new equipment by the viewer with the exception of glasses. This level is said to be HD conventional display compatible (CDC).

Level 2 relates to signals which provide for systems that require a new display but not a new set top box. This level is said to be conventional HD frame compatible (CFC). The 3DTV signal appears as a single HD signal to the set top box, which passes it through to the (new) display, where it is decoded and displayed as left and right pictures. If a 2D service of the same channel or programme is needed, it can in principle be provided as a conventional HD signal simulcast, provided there is sufficient spectrum. The left-eye and right-eye signals do not have the same "spectral occupancy" as conventional HD signals – some has to be sacrificed.

Levels 3 and 4 relate to systems which require a new set top box and a new display, but which offer a normal HD spectral occupancy left-eye and right-eye 3D service. Level 3 is said to be frame–compatible compatible (FCC), because it is an extension of Level 2. Level 4 is said to be conventional HD service compatible (CSC) because an existing 2D set top box will find, in the incoming multiplex, a conventional 2D HD signal which it can pass to a conventional display as a 2D picture.

3.3 Matrix points

Level 1/first-generation profile

The generic receiver here, for which the signal is intended, is a conventional HDTV receiver. The signals transmitted are based on a wavelength division multiplex and matrixing of the left-eye and right-eye signals, and a choice of complementary primary colour separation. For example, the "ColorCode" system has been broadcast in Europe and North America using Red/Green in one eye, and Blue for the other eye. Other sets trialled have been Red vs. Green/Blue, or Green vs. Red/Blue. The exact matrixing and choice of complementary colours can be left to market developments because a conventional receiver and 3-primary display is used, though in the light of experience ITU-R may be able to report on options.

Level 2/first-generation profile

The generic set top box here for which the signal is intended is a conventional HDTV set top box. But the 3D display needed is new and must have the capability to interpret an HD frame as left-eye and right-eye pictures. There are alternative ways to arrange the left-eye and right-eye signals to appear to the STB to be a single frame. The three principal methods (which involve sub-sampling) are the side by side (SbS), the over and under method (OaU), and the interleaved sample (IS), checkerboard or Quincunx method (of which there are variants). BSkyB in the United Kingdom uses the SbS method (2x1080i/960). It would be very valuable to the public to identify a single CFC method for broadcasting. At minimum, a common method of signalling the format is needed, such as has been developed by DVB.

This matrix point may be of particular value to broadcasters who manage a large existing population of set top boxes which must not be disenfranchised by the 3DTV broadcasts, and for whom additional delivery channels are available which can be used for 3DTV.

Simultaneous delivery of a 2D version of the same programme, if needed, requires a simulcast of a conventional HDTV signal.

For this and other Levels, the issue of "creative compatibility" of a 3DTV signal and a 2D TV version needs to be considered.

Level 3/first-generation profile

The generic set top box (or IRD) for which the signal is intended here is a new set top box which is able to decode a Level 2, Frame Compatible image, and also decode a resolution enhancement layer, using for example, MPEG SVC (scalable video coding), yielding normal spectral occupancy L and R HD images for output to the display. This approach would allow existing Level 2 transmissions to be compatibly improved to normal HD spectral occupancy, with the improvement becoming available by replacing the population of conventional set top boxes with the new set top boxes that include a Level 5 H.264 decoder; interlace content would not be supported. Note that unless all set top boxes are replaced, it could still be necessary to simulcast a 2D version of the programme for the 2D audience. Set top boxes (or IRDs) for this level would decode Levels 1 and 2 also.

Level 4/first-generation profile

The generic set top box here is also a new set top box (or IRD) which is able to decode an MPEG MVC signal conforming to the ISO/IEC JTC1 MPEG specification. The signal is arranged so that a conventional set top sees a single 2D HD signal which can be passed to a conventional display as a 2D service. New set top boxes (or integrated receiver/displays) recognize the additional information in order to decode a second view and provide two output signals L and R, to the display. Set top boxes for Level 4/first-generation profile include capability for Level 2 decoding (but, depending on market conditions, not complete Level 3 decoding including extension).

This matrix point may be particularly valuable to operators of terrestrial broadcasting services, where channels are scarce, and where it is necessary to provide both a 3D and 2D service from the same channel.

Level 4/second-generation profile

The generic receiver here for which the signal is intended is also a new set top box which is able to decode the 2D HD plus depth format as specified by the IEC/ISO JTC1 MPEG specification. The display is normally a multiview auto- stereoscopic display. Such set top boxes would also decode Levels 1, 2, and 4 of the first-generation profile.

Other matrix points are left empty for the time being.

4 First-generation 3DTV

It is not currently envisaged that a complete transition from 2D to 3DTV broadcasting will take place in the foreseeable future.

Rather, there is a need to first properly assess the viability of first-generation 3DTV broadcasting. This might perhaps take the form of various 3DTV programme content being made available to the public in a limited ad hoc manner, perhaps just a few hours per week. This could align with other research that is required, such as on the possible effects of eye strain and to assess whether there is acceptance of prolonged stereoscopic viewing. This may be considered as a test phase.

The business models are not the same for pay television and for free-to-air broadcasters, and so the acceptable solutions for first-generation 3DTV broadcasting are anticipated to be different, as explained in § 3 above.

Two variants of first-generation systems may therefore be required for use in different situations: where a service is to be delivered only to viewers with 3D displays; secondly, where the primary audience continues to be viewers receiving an existing 2D service, and it is wished to make use of the same transmission channel to deliver at least some programmes in 3D.

Two techniques are available to satisfy the above conditions:

1 A “frame-based” approach: package the left and right images into an existing HDTV frame.

There are several possible permutations of placement of the left-eye and right-eye images within the frame:

- side-by-side;
- over/under;
- line/column interleaved;
- checkerboard/quincunx.

There is also the potential to add layering techniques to restore the resolution that would otherwise be lost by the placement of two images within a single frame.

A frame-based service would not be directly viewable by existing 2D viewers.

For a multichannel pay television operator, the priority is likely to be to exploit the existing infrastructure in order to deliver 3DTV content to a group of subscribers. Indeed, such an operator could be in a position to do so without impact on services already being delivered to viewers. In this situation, a frame-based solution may be attractive.

A free-to-air operator with access to only limited transmission capacity might require to continue to use existing transmission channels to reach the general 2D audience. In this situation, a frame-based approach would not be suitable.

2 A 2D compatible method

This approach requires that additional information be conveyed in order to reconstruct the second image for suitably equipped 3D receivers.

There are several possibilities for making available the additional information needed to reconstruct the second image:

– simulcast;
– 2D + "delta" (data coded to represent the difference information between left-eye and right-eye images);
– 2D + DOT (data to represent depth, occlusion and transparency information).

  A "2D + depth" coding scheme could allow multiple views to be generated for presentation on autostereoscopic displays.

The 2D compatible approach allows existing viewers to continue to watch a 2D service. Those viewers wishing to receive 3DTV transmissions would need specially equipped receivers.

As an example, Korea's terrestrial broadcasters, KBS, MBC (Munhwa Broadcasting Corp.), SBS and EBS (Educational Broadcasting System), have prepared for 3D trial broadcasting from October 2010 using dual stream coding (left image with MPEG-2, right image with AVC/H.264) at a resolution of 1 920 × 1 080 interlaced 30 fps. Unlike some countries that have already tested 3D TV broadcasting, they will offer the service through terrestrial networks. Furthermore, Korea will be the first country in the world to offer a full HD 3D broadcasting service. In addition, cable broadcasters CJHelloVision and HCN and Korea Digital Satellite Broadcasting, will also take part in the 3D trial broadcasting service.

5 Future generations of 3DTV

Advanced forms of autostereoscopic display in conjunction with multiple camera systems are under study with the intention of allowing viewers to set their preferred viewpoint and to change it continuously in a range determined by the number of cameras and their allocation, for example so-called "free viewpoint television", see Annex 3. This approach can retain backwards compatibility with the displays used for first-generation 3DTV.

There are also studies on possible new forms of "object wave recording" that could allow three-dimensional television images to be presented in a way that represents viewing the physical light in a virtual a space, perhaps using an advanced "integral" method or a holographic system. Such schemes are in the research phase. These studies are to be encouraged, as they promise to lead to the eventual realization of the ultimate goal of presenting images to viewers that are virtually indistinguishable from natural real-world surroundings. To achieve this, new types of advanced volumetric display will be required. It is currently uncertain when this technology might become available: it is likely to be many years in the future.

6 Expected bandwidth requirements for a first-generation system

In the case of a first-generation "2D compatible" system of broadcasting, some additional bit rate will certainly be required. In the extreme, 100% extra would be required for a second simulcast video channel. In practice this would be likely to be somewhat lower using a supplementary data stream for reconstruction of the second video image.

In the case of a "frame compatible" system, if it is accepted that the L and R images contain less spatial resolution than for a 2D system, then in principle no extra bit rate is required compared to the transmission of a normal 2D service. In practice, it is understood that operators plan to use broadcast bit-rates which are at the high end of current practice. It may nevertheless be argued that because this approach does not provide a 2D-compatible service, a completely new transmission channel is required, i.e. 100% extra capacity. However, in the circumstances of a multichannel operator this might not necessarily be a constraint.

It is currently unclear what the quality differences would be between these approaches, and there will inevitably be a trade-off between bit-rate and quality. Independent testing would be desirable.

The human visual perception can be exploited to reduce bandwidth requirements. For example:

– filtering (blurring) in one eye (switching on scene cuts);

– asymmetrical coding.

The 2D + depth approach offers the prospect of considerable bit-rate saving. However, a cost-effective method for depth map creation is not easy to obtain and is still an active area of research.

In the case of more advanced multiview schemes, multiview coding requires multiple synchronized video signals to show the same scene from different viewpoints. This leads to large amounts of data, but typically a larger amount of inter-view statistical dependencies than for stereo.

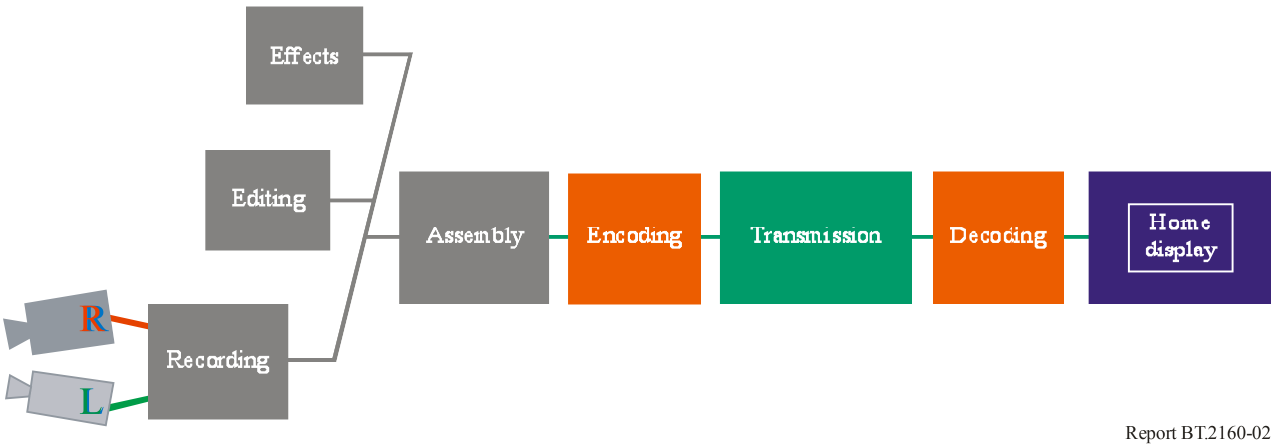
Last, but not least, independent testing using a standardized testing methodology is needed in order to accurately quantify how much extra bit rate would be needed, using a range of representative 3DTV source material.

7 The 3DTV broadcasting chain

The end-to-end broadcasting chain from image source, programme production, delivery and display may be illustrated as follows:

FIGURE 2

The broadcasting chain (indicative)



Report BT.2160-02

The implications for the following individual elements of the broadcasting chain should be considered:

7.1 Image source methods

There are three main approaches to sourcing 3D programme material in use today. These are: stereo camera, CGI, and conversion from 2D video.

Most 3D video captured presently use stereo camera rigs. Some test footage has been captured using stereo cameras coupled with a rangefinder. Rangefinders are usually laser or infrared-based and attempt to provide depth maps for a given scene. The depth maps are prone to numerous errors due to a number of issues, such as poor accuracy, speculars, translucent objects, transparent objects and reflections. Another capture method that has had some testing is multi-camera rigs. These have a large number of cameras that provide a number of views. This method works well for capturing several views. However, the complicated rigs plus the large amounts of data currently prohibit widespread use of multi-camera rigs.

Computer generated content is typically considered the easiest method of stereo generation. The rendering system can either render one or more related views depending on the application. In addition, the Z buffer, which represents the distance to the screen of various objects, can be exported as a depth map. In either case, computer generated data can be used for stereoscopic production or for multiview production.

Finally, 3D video can be created by taking conventional 2D video and adding depth information. The normal process is to deconstruct the 2D image into a series of objects (also known as segmentation), assigning relative depth to each object, then filling in occluded areas. Human visual perception can also be exploited in the processes for converting from 2D to 3D. The creation of a depth map from a 2D allows for the creation of multiple views, through a rendering process that incorporates techniques of covering disoccluded regions.

7.2 Characteristics of signals in the studio

Without coding or compression, the baseband required for a two-channel 3DTV system, with HD resolution for each eye, is twice that required for a HDTV system. However, the actual requirements will depend on the format of the signals of the 3DTV system.

– How much information is involved?

– Can signals be handled by existing equipment and interfaces?

– Would new interfaces be required?

The answers to the above questions can be expected to vary according to the form of the 3DTV system.

Some form of metadata for first-generation systems are required to ensure that the left-eye and right-eye views are correctly identified. This may be based on either explicit or implicit information. For example, in the side-by-side format, whether the image on the left consists of the left-eye or right-eye view, and the sampling structure used, has to be known. Synchronization of the left-eye and right-eye views is also needed to ensure that there are no errors in timing, such as with the above-below format. Some of these signals might be able to be handled by existing equipment and interfaces but others might not.

Also, control signals are required for active eyewear that has to synchronize its operation with the view that is being displayed on the screen.

7.3 Programme production

Equipment is required that must handle recording, editing, effects, and postproduction.

The effect of the introduction on first-generation 3DTV on existing Recommendations that apply in the studio production environment will need to be considered.

Suitable provision will need to be made for monitoring the quality of the 3DTV at the point of origination and at appropriate points in the production chain.

One UK satellite broadcaster's technical guidelines, the prime objectives of which are to deliver content of both a high technical quality and of high production values may be found at http://introducingsky3d.sky.com/a/bskyb-3d-tech-spec/

http://www.sky.com/shop/tv/3d/producing3d.

In the UK, the BBC has issued interim delivery requirements for programmes made in stereoscopic 3D. See: http://www.bbc.co.uk/guidelines/dq/pdf/tv/tv_delivery_to_network_programmes_v1.2-2011.pdf

Further study is required.

7.4 Emission

3DTV signals may need to be encoded in ways that are appropriate to their transmission within the existing 6/7/8 MHz terrestrial transmission channels, and also by existing broadcast satellite services.

Different techniques are likely to be required that are appropriate to each of these situations, and according to the requirements of the broadcaster as indicated in §§ 4 and 5.

7.5 Display

There are known to be three fundamental approaches:

– viewer wears glasses;

– without glasses (auto-stereoscopic);

– headmounted display.

With a headmounted display, the left and right eyes are presented with the left-eye and right-eye images of a stereo pair. This may be appropriate for video games, but is unlikely to be appropriate for viewing of broadcast television. It is an individual viewing experience and is not suitable for collective (e.g. family) viewing of broadcast television.

Within these broad categories, various approaches may be possible. In many cases, 3DTV presentation relies on some form of eyewear or headgear that the viewer must wear in order to discriminate between left-eye and right-eye images:

– Anaglyph: a stereoscopic effect can be obtained by displaying images in which the presentation screen simultaneously displays two differently-filtered coloured images, (typically red for the right-eye image and cyan for the left-eye image). These are viewed through correspondingly-coloured glasses. One difficulty with this solution is that the viewer may feel compelled to remove the coloured glasses when looking away from the presentation screen. In addition, the programme presentation will necessarily provide an inferior colour rendition.

– Polarized glasses: this solution makes use of cross-polarizations for the right-eye and the left-eye images of a stereo pair; the images are watched through correspondingly cross-polarized glasses. One solution to display such cross-polarized image uses a "tiled" display of alternating tiles for the first and second image of a stereo pair. The tiled display is covered by an identically tiled polarized mask, with alternating tiles being cross-polarized. When viewed through cross-polarized glasses, separate views will be presented to the left and right eyes of the viewer. One problem with this solution is that the presentation of stereo images at HDTV resolution requires a more expensive display providing at least twice the horizontal resolution of HDTV.

– Shuttered glasses: the two images of a stereo pair are time-interleaved on the screen, and viewed through special glasses in which the left and right eye lenses are shuttered in turn, following the switching cycle of the left and right images on the screen.

While it seems inevitable that, at least to begin with (and possibly for many years) 3DTV viewing would require that viewers to wear glasses, first-generation 3DTV broadcasts could nevertheless continue to be viewed on more advanced improved forms of autostereoscopic display as the technology progresses.

8 Production grammar

Poor quality stereoscopic television could “poison the water” for everyone. There is a risk that 3DTV becomes associated with eye strain if stereoscopic content is poorly realized – whether due to inappropriate production grammar or due to inadequate technology for delivery. This has happened before in the cinema in the 30s, 50s, and 80s.

The production grammar of 3D often differs to 2D productions. Special care has to be taken in order to achieve a good 3D viewing experience. This can lead to some compromises for the 2D viewer. In some cases, a production might be optimized for 3D, with no intention that the 3DTV version be used for conventional standard- or high-definition television presentation.

It is understood that various recent trial 3DTV productions have provided useful learning experiences, and it is expected that further knowledge will be gained through ongoing trial productions and from services that have recently become available to the public. Live 3DTV production presents particular challenges. Live production nevertheless forms a regular part of the schedule of the recently introduced 3D service from United Kingdom pay television operator BSkyB.

9 The viewing environment

The effect of the viewing environment is fundamental on the perception of depth and to the quality of the overall viewing experience. The following situations should be considered:

– the studio environment;

– the home environment.

In particular, in conjunction with viewing distance, picture size and subtended viewing angle play a role in the three-dimensional effect as perceived by the viewer. This might have implications on the way in which 3DTV should be produced and displayed.

10 Principles for comfortable viewing of stereoscopic three-dimensional images

10.1 Composite factors in perception of stereoscopic 3D images

A stereoscopic 3D system expresses depth by presenting video that has parallax with respect to the left and right eyes of the viewer. Perception depends on programme production techniques, display devices, 3D glasses, viewing conditions, and viewer characteristics. Consequently, visual fatigue associated with the viewing of stereoscopic 3D programmes will be affected by composite factors consisting of programme production techniques, display devices, 3D glasses, viewing conditions and viewer characteristics.

NOTE – Section 5 of Annex 4 describes the spatial distortion prediction system for 3DTV that calculates the spatial distortion of a reproduced stereoscopic image and predicts the extent of the puppet-theatre and cardboard effects, excessive binocular parallax, and excessive parallax distribution.

10.2 Measures to enable comfortable viewing of stereoscopic three-dimensional images

All parties concerned with stereoscopic 3DTV systems should take the above characteristics of stereoscopic 3D systems into account when manufacturing equipment, producing programmes, displaying video, and viewing video programmes. Solely regulating the amount of parallax in 3DTV programmes is not reasonable.

Due to the complexity of the end-to-end broadcast chain that involves many organizations and technologies, from capture, through production, mastering, broadcast, and reception to display, no single organization has end-to-end control over this effect.

10.2.1 Programme production

It may be useful to identify measures to help avoid the inadvertent creation of materials for transmission on broadcast television that would likely induce visual fatigue and other possible health hazards. Measures should be proportionate to the risks and should not place undue burdens on broadcasting organizations or programme producers. The impact of measures on broadcasters or programme producers may vary with their programme genres. For example, programme production is often beyond the control of the broadcaster in some live programming, such as news.

Broadcasting organizations should be encouraged to raise awareness among programme producers of the risks of creating stereoscopic television image content that may induce visual fatigue and create other possible health hazards in viewers of stereoscopic television broadcasts.

Producers of 3D programmes need to understand the characteristics of stereoscopic 3D images and the various effects of 3D video techniques.

10.2.2 Viewing environments and display devices

Viewing environments and display devices, which can affect the likelihood of problems, may differ between households, reflecting the style of living. Nevertheless, viewers should be well informed of satisfactory conditions for viewing stereoscopic 3D images. Example notifications given to viewers in Japan are described in Annex 7.

11 Psychophysical aspects of viewing stereoscopic images

11.1 Psychophysical aspects

Before attempting to implement new broadcast schemes, it is necessary to gain a full understanding of the results of psychophysical studies in order to understand the effects to which the viewer is subjected and the performance that is required of the main equipment in these systems. There are a number of issues to be studied before the effects of viewing three-dimensional images on human perception and visual functions can be fully understood.

Section 1 of Annex 4 identifies some key study items on the psychophysical aspects of stereoscopic television systems. It also includes the results of studies related to the naturalness and unnaturalness of stereoscopic video, the evaluation of visual comfort based on an analysis of parallax distributions within certain frames, and the visual fatigue by viewing stereoscopic video.

An important problem of current stereoscopic television, which is common to all the approaches currently implemented, is that they present stereoscopic images on a single surface (the display screen), giving rise to a potential conflict between “vergence” (the eye movement to point both eyes to the same point on the screen) and “accommodation” (the action by which the “lens” in the viewer’s eye focuses on that point). It has been documented in medical literature that this conflict can cause viewer's discomfort, eye fatigue, headache and possibly other health hazards notably if the viewing continues for an extended period of time[3,4].

Such effects have been noticed in some recent trial broadcasts. For example, one of the terrestrial broadcasters in Korea, SBS (Seoul Broadcasting System), broadcast South Africa World Cup Soccer Games in 3D from 11th June 2010. In a survey of nearly 100 viewers, 75% expressed satisfaction with the trial 3DTV broadcasting service. However, the survey showed that 30% of the viewer felt dizziness, double image and eye fatigue.

Factors that affect 3D viewing comfort also include inter-pupillary distance, intra-scene disparity range, and the speed of depth change of objects in the scene. In addition, rapid cuts between shots of differing depths and changing depths with zoom or pans are known to cause viewer discomfort. Some of these techniques are widely used in 2D production but might cause discomfort when viewed in 3D. Due to these factors, 3D production techniques tend to create 2D video that might be considered by 2D viewers as boring. This is the reason that many 3D productions to date have been different from the 2D productions of the same event or release. It is widely known that current 3D movie releases are editorially different from the 2D releases. Parallax is affected by the programme production technique, display device, viewing conditions, and viewer characteristics (such as inter–pupil distance). Accordingly, all parties concerned with stereoscopic 3D systems should take this characteristic of stereoscopic 3D systems into account when manufacturing equipment, producing programmes, displaying video, and viewing video programmes.

11.1.1 Geometrical relationships and naturalness

The reproduction of depth information is essential for people to gain a sense of three-dimensionality from a stereoscopic image. Depth distortion in the stereoscopic image can create an unnatural impression when the image is viewed.

In stereoscopic imaging, the object is imaged using two cameras. The arrangement of these two cameras can be classified into parallel configurations where the optical axes of the two cameras are parallel with each other, and intersecting "toed-in" configurations where the two optical axes are made to intersect. When using a parallel configuration, a stereoscopic image with no spatial distortion is obtained when the gap between the cameras is set equal to the gap between the pupils, the horizontal offset of the left and right images projected on the screen is equal to the gap between the pupils, and the camera's angle of view is equal to the expected viewing angle of the display. In such cases, the image is said to be viewed under orthostereoscopic (distortion-free) conditions. When actual programme production and viewing conditions are taken into consideration, it is difficult to ensure that distortion-free conditions are always satisfied. If these conditions are not met, depending on the conditions, the spatial distortion of the stereoscopic image can cause

---

[3] See for instance: K. Ukai & P.A. Howarth "Visual fatigue caused by viewing stereoscopic motion images: background, theories and observations" – Elsevier B:V., 2007, which states, *inter alia*, "Viewers should be careful to avoid viewing stereoscopic images for extended durations because visual fatigue might be accumulated.

They should be ready to stop immediately if fusion difficulties are experienced.

Hardware/software manufacturer should avoid unnatural image presentations, such as images that diverge further than infinity, large binocular disparity in the central visual field or around the objects that are the centre of the viewer's attention, difference of size and colour, unequal distortion between binocular images since they may cause fusion difficulty".

[4] Masaki Emoto, Takahiro Niida, and Fumio Okano - "Repeated Vergence Adaptation Causes the Decline of Visual Functions in Watching Stereoscopic Television", Journal of Display Technology, Vol. 1, No. 2, December 2005.

unnatural effects such as the “puppet theatre” effect and “cardboard” effect. The puppet theatre effect is a phenomenon wherein the stereoscopic images of foreground objects appear unnaturally small. The cardboard effect is a phenomenon wherein the stereoscopic image of an object appears unnaturally thin.

Section 2 of Annex 4 presents a geometric analysis of reproduced stereoscopic image spaces and discusses the results and their relationship to the distortion of reproduced stereoscopic image spaces. The results of subjective evaluation tests that support these findings are also shown. The discussion relates to how the reproduced stereoscopic image space is affected by parameters such as the camera configuration (parallel or toed-in), display screen size, and viewing distance.

11.1.2 Visual comfort and discomfort in viewing stereoscopic images

Finding a way to make the visual comfort of stereoscopic images a measurable physical factor is arguably one of the key issues in stereoscopic imaging research. Stereoscopic images convey depth information to the viewer by making use of the parallax between the images presented to the left and right eyes. If we could ascertain how the magnitude and distribution characteristics of this parallax relate to the visual comfort of the image, this information would be very useful for the production of stereoscopic images.

How these parameters relate to the subjective visual comfort of stereoscopic images was studied by focusing on the average and range of the parallax distributions. It was shown that they both have a correlation with the visual comfort of stereoscopic images and that the range of parallax distributions in stereoscopic images appraised as visually comfortable was almost 60 pixels in HDTV image. It was also suggested that stereoscopic images tend to become more visually comfortable when the average value of the parallax distribution approaches zero (i.e. at apparent positions closer to the display screen).

In stereo 3D systems, a binocular 3D image is formed by presenting each of these images to the respective right and left eyes. If discrepancies arise between these two images due to the systems used for production, storage, transmission or display, they can cause psychophysical stress, and in some cases 3D viewing can fail. For example, when shooting and displaying stereoscopic 3DTV programmes, there can be geometrical distortions, such as size inconsistency, vertical shift, and rotation error, between left and right images. It is desirable that these geometrical distortions should be suppressed. Stereoscopic image cross-talk, in which the images “leak” and can be partially seen by the opposite eye, can also result in discomfort for the viewer. Detection and tolerance limits of evaluating visual discomfort in terms of cross-talk were reported to be highly dependent on image content and display contrast, and cross talk must be reduced on high contrast displays.

Section 3 of Annex 4 presents some results of subjective evaluation tests with regard to visual comfort in viewing stereoscopic images. The results indicate that stereoscopic image having an excessive range of parallax distribution can be evaluated as uncomfortable to view. The research results on visual discomfort caused by discrepancies between left and right images are also shown. The results indicate the detection and tolerance limit of discrepancies with regard to visual discomfort in viewing stereoscopic images.

11.1.3 Visual fatigue in viewing stereoscopic images

One of the major factors of visual fatigue caused by viewing stereoscopic images is the difficulty in fusing left and right retinal images with large binocular parallax, which lead to increased viewer fusion effort. Fusion effort is based on two factors: the principle of stereoscopic display (defined by horizontal binocular parallax, inevitable in stereoscopic systems), and issues involving hardware (leading to differences between views of left and right images).

Another important aspect of stereo 3D systems is that a dissociation of vergence and accommodation can be a major factor of visual fatigue in viewing stereoscopic images. This is because a difference in visual functions between viewing real objects and viewing stereoscopic images. The vergence point is positioned within the depth of field when viewing a real object. On the other hand, the vergence point is sometimes outside the depth of field when binocular parallax is large in viewing stereoscopic images. Temporal discontinuous changes in dissociation can also lead to visual fatigue.

Section 4 of Annex 4 presents some experimental results of subjective evaluation with regard to visual fatigue in viewing stereoscopic images. The results indicate that inconsistencies between vergence and accommodation can cause visual fatigue.

11.2 Examples of safety guidelines

In 2010, a liaison statement (see Document 6/316) was sent to the WHO requesting them for information on potential impact on health of 3D. They recently replied after a reminder. They could not give any direct information from their own files as they presently do not have a project on this topic. Guidance has nevertheless been made available by some national bodies.

11.2.1 Korea (Republic of)

A 3DTV Project Group (PG806) has been established in Korea with the aim of development of 3DTV broadcasting specification and viewing safety guideline. Working Group WG8062 is focusing on the development of a 3DTV viewing safety guideline for display, contents, viewing condition, and viewer parameters.

TTA published "3DTV Broadcasting Safety Guideline" in December 2010, see Annex 5. Its purpose is to present a way to reduce the visual fatigue and conflict of viewing stereoscopic contents, and to promote related 3D applied industries by reducing potential risk factors for the viewing stereoscopic contents. This guideline is intended to present adequate circumstances of 3D viewing, notes for viewer, suitable use conditions of contents and display guideline for safe viewing of 3D broadcasting service. It is planned to update this guideline reflecting a result of clinical research for 3DTV viewing.

Recent actions taken by the Italian Health Ministry, related to the use of 3D spectacles by the public attending cinema presentations of 3D movies, are described in Annex 6.

Factors that affect 3D viewing comfort also include inter-pupillary distance, intra-scene disparity range, and the speed of depth change of objects in the scene. In addition, rapid cuts between shots of differing depths and changing depths with zoom or pans are known to cause viewer discomfort. Some of these techniques are widely used in 2D production but might cause discomfort when viewed in 3D. Due to these factors, 3D production techniques tend to create 2D video that might be considered by 2D viewers as boring. This is the reason 3D productions to date have been different from the 2D productions of the same event or release. It is widely known that current 3D movie releases are editorially different the 2D releases.

12 Assessment methodology

Although a method for subjective assessment of image quality and depth quality is provided by Recommendation ITU-R BT.1438 – Subjective assessment of stereoscopic television pictures, the type and visibility of artefacts peculiar to stereoscopic images have yet to be systematically identified and studied. Furthermore, the various methodologies and formats have to be taken into consideration.

The development of an appropriate assessment methodology, in conjunction with a common set of reference source material is of the utmost importance for evaluating 3DTV systems. It is understood that PSNR results might not be indicative of the effect of artefacts, and that new metrics will need to be considered. Major issues concern the identification of the factors that contribute to viewing discomfort and the development of proper metrics for measuring levels of discomfort. It is especially urgent to not only seek a metric for the measurement of viewing comfort, as this is a major concern for most users and providers alike, but also to seek a methodology for testing viewing comfort for both short-term and long-term viewing.

Visual comfort, image quality and depth quality are major perceptual dimensions that both users and programme providers are interested in. However, the value of routine testing of other perceptual dimensions, such as “presence”, “sensation of reality” and “naturalness”, should also be investigated. It is likely that new metrics and methodology are required. Methodology is also required to compare the performance of various approaches to the transmission of 3DTV signals and effects of bandwidth reduction.

13 User requirements

These are currently not fully understood.

At its May 2009 meeting, WP 6C decided to carry out a survey on the aspirations of the ITU Membership on 3DTV broadcasting. The survey was carried out between July 2009 and October 2009. All those who responded considered that there is a need to discuss with standards bodies, such as the IEC, the provision of minimum requirements for 3DTV receivers which match a future 3DTV broadcast system. In addition all responders considered it very important/essential for a 3DTV system to have the same format as packaged media (e.g. HDTV capacity discs).

14 Performance requirements

The overall performance requirements need to be identified in sufficient detail in order to orient the choice of the appropriate technologies for a new 3DTV system.

A preliminary list of possible requirements is listed in Annex 9. This has been gleaned from contributions made to WP 6C since October 2008.

It is hoped that future contributions will provide clarifications and/or additional factors that should be considered.

15 Organizations with initiatives in 3DTV

A wide range of research, standardization, and trade associations are currently active in investigating aspects of 3DTV. A non-exhaustive list is attached in Annex 1.

16 Conclusions

Without an orderly approach to the standardization of 3DTV broadcasting systems, even for an initial test phase, various *de facto* standards will become established. There is a risk that subsequent implementation of 3DTV broadcasting could become more difficult.

Furthermore, actions likely to be taken by the gaming and optical media (Blu-ray) industries could have a significant impact on the capabilities of widely deployed consumer equipment.

It is also not known what the consequences might be of decisions on the future 3D-capabilities of interfaces to displays if these are taken in the absence of agreed requirements for 3DTV broadcasting systems.

It is anticipated that guidance will be desirable covering the following:

– quality assessment methods for 3DTV systems;
– reference 3DTV source materials for use in subjective tests;
– requirements for the broadcasting chain;
– requirements for production and production grammar;
– psychophysical aspects related to viewing of stereoscopic images;
– requirements first-generation 3DTV systems.

In addition, an important issue for further study is an understanding of bit-rate requirements for first-generation 3DTV broadcasting systems, for both the frame-based and compatible 2D approaches.

Referring to the matrix of signal formats described in Fig. 1, the most critical matrix points that might need to be standardized are the first-generation Levels 2 and 4 points to the maximum extent possible, but certainly regarding signalling.

Another critical issue is to try to align the matrix with the formats used for packaged media.

Further contributions to WP 6C are invited on the above and related topics.

Annex 1

Organizations with current initiatives in 3DTV

1 ISO/IEC JTC1/SC29/WG11

In July 2009 it was planned to finalize the specification of carriage of MVC over MPEG-2 systems, as well as extensions to the file format specifications to accommodate multiview video.

Work is also proposed to begin on a new 3D video (3DV) format that aims to support advanced stereoscopic display processing and auto-stereoscopic displays.

An amendment to ISO/IEC 14496-10 includes a spatially interleaved frame supplemental enhancement information (SEI) message to signal the type of interleaving in a frame-based scheme.

2 ITU-T Study Group 9

ITU-T SG 9 has recently initiated work items to develop draft new Recommendations on subjective assessment methods for 3D video quality and on display requirements for 3D video quality assessment. In addition, studies are being progressed on scalable view-range representation for free viewpoint television (FTV).

3 ITU-T Study Group 16

The multiview coding extension of Recommendation ITU-T H.264 | ISO/IEC 14496-10 MPEG-4 AVC has proceeded to AAP Consent under ITU-T Recommendation A.8 approval process.

4 3DTV – Network of Excellence

See: www.3dtv-research.org.

5 3D4You – Content generation and delivery for 3D television

3D4You was a project funded by the European Union under the information and communication technologies (ICT) Work Programme 2007-2008, a thematic priority for research and development under the specific programme "Cooperation" of the Seventh Framework Programme (2007-2013). Project website: http://www.3d4you.eu/.

6 SMPTE

The activities of the Society of Motion Picture & Television Engineers (SMPTE) include standardization work related to stereoscopic 3DTV in the production environment. SMPTE's work is distributed among its various Technology Committees, Working Groups and Ad Hoc Groups

One such activity has been completed and has resulted in a published standard (SMPTE ST 292-2:2011 Dual 1.5 Gb/s Serial Digital Interface for Stereoscopic Image Transport) that defines a method of transporting stereoscopic images using two streams of 1.5 Gbit/s in conjunction with the means to identify each stream.

For each current activity related to stereoscopic 3DTV, the table below identifies the responsible SMPTE group, the scope of work, and its status as of September 2011. See also: https://www.smpte.org/.

Overview of SMPTE standardization activities related to stereoscopic 3DTV (September 2011)

Standards Committee Group	Name	Scope	Status
Technology Committee 35PM, Working Group on 3D Home Master	ST 2061 Stereoscopic Distribution Master	The Stereoscopic Distribution Master (formerly known as the '3D Home Master') is intended to provide a standardized means for interchange of 3D content amongst mastering facilities, and between a mastering facility and the ingest facility of a distribution system. The Stereoscopic Distribution Master may feed various distribution outlets for 3D content to the home, including (but not limited to): mobile, Blu-ray/DVD, streaming, terrestrial, and cable/satellite broadcast. The document includes a Glossary, and covers Image Structure, Subtitles, Captions and Graphical Overlays, and Metadata	FCD Final Committee Draft

(*continued*)

Standards Committee Group	Name	Scope	Status
Technology Committee 10E Essence: AHG	ST 2066 Disparity Map Representation for Stereoscopic 3D	Identify requirements for a data representation of disparity maps relevant for production, post-production, and distribution of 3D content.	Work in progress
	ST 2068 Frame Compatible	This document categorizes and enumerates the various methods to transport a pair of stereoscopic frames within single video frames. Unique identification codes are assigned to these methods in order to facilitate interchange of 3D Frame Compatible signals. The information contained can be used to confirm that proper encoder/decoder pairs are used. It can also set the H.264/MPEG-4 AVC Frame Packing Arrangement SEI message prior to delivery to end user terminal devices. One purpose of this document is to carry the frame packing information necessary for the MPEG-4 AVC Frame Packing Arrangement SEI message. A second purpose is to carry information describing the frame packing arrangement that may be used in the production workflow. The document will not enumerate the filtering used prior to sub-sampling, but optional suffix coding may be provided in some implementations in order to identify unique filtering on the basis of specific implementations.	Work in progress
Technology Committee 32NF Networks and Infrastructure	SMPTE ST 2063. Stereoscopic 3D Full Resolution Contribution Link – MPEG-2 TS	This document specifies how a stereoscopic 3D video system based on the MPEG-2 Transport Stream (TS) that is codec agnostic (i.e., any codec for which there are defined methods for transport via MPEG-2 TS is permitted) performs coding, multiplexing, and decoding. It defines constraints for the input image pair, the bit stream, the multiplexing, timing synchronization, and signalling, as well as for the video coding and decoder behaviour. The input image pair must have the same image structure (horizontal and vertical pixel count, scanning system, colorimetry, and frame rate) and be coincident in time.	Final Committee Draft

(*end*)

	ST 292-2 Dual 1.5 Gb/s Serial Digital Interface for Stereoscopic Image Transport	This standard defines a means of transporting stereoscopic images (Left eye and Right eye images) using an interface consisting of two links based on the SMPTE ST 292-1 data structure. The Left eye images are carried on one link of the interface and the Right eye images are carried on the other link. The stereoscopic image formats to be transported using this standard are the 4:2:2 10 bit image formats defined by SMPTE ST274, ST2048-2 and ST296, which can be transported by a single SMPTE ST292-1 serial interface. Audio and other associated ancillary data may also be transported. This standard also defines a payload identifier.	Published standard
	ST 425-2 Source Image Format and Ancillary Data Mapping for Stereoscopic Image Formats on a single-link 3Gb/s Serial Interface	This standard defines a means of transporting a stereoscopic image pair consisting of a Left Eye and Right Eye image (Le and Re) using an interface consisting of a single 3Gb/s (nominal) link. The stereoscopic image formats to be transported using this standard are those 4:2:2 10 bit image formats having a sampling frequency of 74.25 MHz, or 74.25/1.001 MHz. Audio and other associated ancillary data may also be transported. This standard also defines a payload identifier. It is not necessary for implementations to include support for all formats defined in this standard to be compliant. Implementers should indicate supported formats in commercial publications.	Final Committee Draft
	ST 425-4 Dual 3 Gb/s Serial Digital Interface for Stereoscopic Image Transport	This standard defines a means of transporting stereoscopic images (Left eye and Right eye images) using an interface consisting of two streams based on the SMPTE ST 425-1 data structures. The Left eye images are carried on one stream of the interface and the Right eye images are carried on the other stream.	Work in progress
Technology Committee 31FS	3D interleaved in MXF OP 1a	File format to standardize the transport of left and right eye images in frame interleaved MXF files for use in TV acquisition, contribution, distribution, station operations, and archives.	Work in progress

7 The Digital Video Broadcasting Project

Technical work in Digital Video Broadcasting Project (DVB) is driven by commercial requirements. Following completion of a study mission to investigate the possible need for 3D activities, further activity led to the publication of “DVB commercial requirements for DVB-3DTV” (DVB Document A151, July 2010). This addresses “frame compatible” 3DTV services over HD broadcast infrastructures. Work has started on considering a second phase of "2D service compatible" commercial requirements.

On 17 February 2011 the DVB Steering Board approved the DVB-3DTV specification, which has been published as BlueBook A154 "Frame Compatible Plano-Stereoscopic 3DTV", (DVB-3DTV). The specification has been sent to the European Telecommunications Standards Institute (ETSI) for formal standardization.

The specification specifies the delivery system for frame compatible plano-stereoscopic 3DTV services, enabling service providers to utilize their existing HDTV infrastructures to deliver 3DTV services that are compatible with 3DTV capable displays already in the market. This system covers both use cases of a set-top box delivering 3DTV services to a 3DTV capable display device via an HDMI connection, and a 3DTV capable display device receiving 3DTV services directly via a built-in tuner and decoder.

8 The Blu-ray Disc Association (BDA)

See: http://www.blu-raydisc.com.

9 HDMI Licensing, LLC, has announced the release of HDMIspecification 1.4

See: http://www.hdmi.org/index.aspx.

10 Consumer Electronics Association

The Consumer Electronics Association (CEA) has established a 3D Task Force. This is considering interfaces between consumer sources, sinks, repeaters, converters, and glasses. They are also considering what is needed for "3D READY" products. It is proposed to develop standards for 3D glasses, including interface, signalling, setup, control and polarization. A project is being considered to update CEA-861 to carry 3D content.

11 The 3D@Home Consortium

This comprises around 40 members, with the aim of speeding the commercialization of 3D video into homes worldwide.

See: http://www.3dathome.org.

12 Association of Radio Industries and Businesses

The Association of Radio Industries and Businesses (ARIB) has established a working group for researching 3DTV broadcasting in 2008.

See: http://www.arib.or.jp/.

13 Ultra-realistic communications forum

The Ultra-realistic communications forum (URCF) is a forum established by the organizations from industries, government and academies, with the aim of promoting the R&D of ultra-realistic communications.

See: http://www.scat.or.jp/urcf/.

14 3D Consortium

The 3D Consortium was established in 2003 and comprises 47 members from 3D industry. Its main focus is on stereoscopic 3D.

See: http://www.3dc.gr.jp/.

15 Consortium of 3-D image business promotion

The Consortium of 3-D image business promotion was established in 2003 and comprises 49 members.

See: http://www.rittaikyo.jp/.

16 Japanese Ergonomics National Committee

Japanese Ergonomics National Committee (JENC) is in charge of the national preparation for ISO TC159.

See: http://www.ergonomics.jp/.

17 Telecommunications Technology Association

In Korea, Telecommunications Technology Association (TTA) has established a 3DTV Project Group (PG806), with the aim of development of 3DTV broadcasting specification and viewing safety guideline. 3DTV PG consist of two Working Groups; WG8061 for the development of 3DTV broadcasting specification and WG8062 for the development of 3DTV viewing safety guideline.

TTA subsequently published "3DTV Broadcasting Safety Guideline" in December 2010 (see Annex 5).

See: http://www.tta.or.kr/.

18 European Broadcasting Union (EBU)

The EBU (http://www.ebu.ch) has produced a 3D briefing document for senior management. This is reproduced in full in Annex 8.

In addition, EBU Recommendation R 135 "Production & Exchange Formats for 3DTV Programmes" (August 2011) provides interim recommendations for EBU Members who are required to produce, exchange, archive and distribute 3D programmes using 2D infrastructure and transmission technologies, see: http://tech.ebu.ch/docs/r/r135.pdf.

19 MUSCADE

MUSCADE is intending to create major innovations in the fields of production equipment and tools, production, transmission and coding formats allowing technology independent adaptation to any 3D display and transmission of multiview signals while not exceeding double the data rate of monoscopic TV, and robust transmission schemes for 3DTV over all existing and future broadcast channels. MUSCADE is a collaborative project co-funded by the European Commission's Seventh Framework Programme – ICT under the theme "Networked Media and 3D Internet". See: http://www.muscade.eu/index.html.

20 3D VIVANT

The 3D VIVANT project is investigating the generation of a novel true 3D video technology, based on mixed 3D Holoscopic video content capture and associated manipulation, and display technologies. This project, is supported by the European Commission through the Information & Communication Technologies programme. See: http://www.3dvivant.eu/.

Annex 2

Historical background on the development of stereoscopic and 3D television systems

Document 6C/92 describes the present state of three-dimensional (3D) TV broadcasting studies in the Russian Federation:

6C_92.doc

Annex 3

Introduction to free viewpoint television

See Annex 1 to Annex 6 to Document 6C/69:

069N06e.doc

Annex 4

Psychophysical studies on three dimensional television systems

Summary

Before attempting to implement new broadcast schemes, we must gain a full understanding of the results of psychophysical studies in order to understand the effects to which the viewer is subjected and the performance that is required of the main equipment in these systems.

There are a number of issues to be studied before we can fully understand the effects of viewing three dimensional images on human perception and visual functions. For the success of three-dimensional television broadcasting, all parties concerned, including broadcasters, producers, manufacturers, and regulators, should be well informed of the effects.

The psychophysical aspects of viewing stereoscopic images have been extensively studied. This Annex provides some key study items and the study results on the psychophysical aspects of stereoscopic television systems. It also describes the spatial distortion prediction system for 3DTV that calculates the spatial distortion of a reproduced stereoscopic image and predicts unnatural size distortion, excessive binocular parallax, and excessive parallax distribution on the basis of the shooting, display, and viewing conditions.

1 Key items for psychophysical studies

The following sections describe the key items for which further study is encouraged:

1.1 Naturalness and unnaturalness of images

1) Theoretical analysis of spatial reproduction characteristics of images taken by 3D cameras

   It is of fundamental importance to understand precisely how a real space is converted into a stereoscopic image space by a camera. In particular, the reproducibility of a stereoscopic image space should be analysed in terms of different settings of the lens axes of 3D cameras.

2) 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; this is called the "puppet theatre" effect.

3) Depth distortion

   The imaging and display conditions may reduce the reproduction magnification ratio of the depth direction and distort the perception of objects with visually imperceptible thicknesses. This is called the "cardboard" effect.

Section 2 describes the study results on the naturalness and unnaturalness of stereoscopic images.

1.2 Viewing comfort and discomfort

1) Differences in size, verticality, inclination and brightness, and cross-talk

   Viewers may not feel comfortable 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.

2) Psychological factors and the parallax distribution

The fundamental relationship between psychological effects brought about by 3D images and factors related to fatigue should be studied. In particular, “ease of viewing” and “sense of presence” may be key psychological factors. 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 (visual discomfort).

3) Superimpositions within 3D images

With regard to superimpositions in a two-dimensional image, we only have to think about exactly where to display it on the screen. In the case of a stereoscopic image, however, we also need to pay attention to the depth of the superimposition. If we could find a preferred position for superimposition for stereoscopic images, we will be able to use it for actual programme production.

4) Change in parallax distribution during scene changes

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.

Section 3 describes the study results on the viewing comfort and discomfort of stereoscopic images.

1.3 Visual fatigue caused by parallax 3DTV viewing

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

Section 4 describes the study results on the visual fatigue caused by viewing stereoscopic images.

1.4 Individual differences in the stereopsis function

Visual functions vary greatly from person to person, so it is essential to understand that there are individual differences before 3D broadcasts begin. 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.

1.5 Effect on young people

We must bear in mind that young people’s sense of sight changes as they mature. Viewing of stereoscopic images possibly affects their visual functions in ways different from adults. It may be advisable that young children be cautioned about viewing stereoscopic images for extended periods of time.

2 Naturalness and unnaturalness of stereoscopic images − Geometrical analysis of spaces reproduced by stereoscopic images

2.1 Theoretical analysis of reproduced spaces

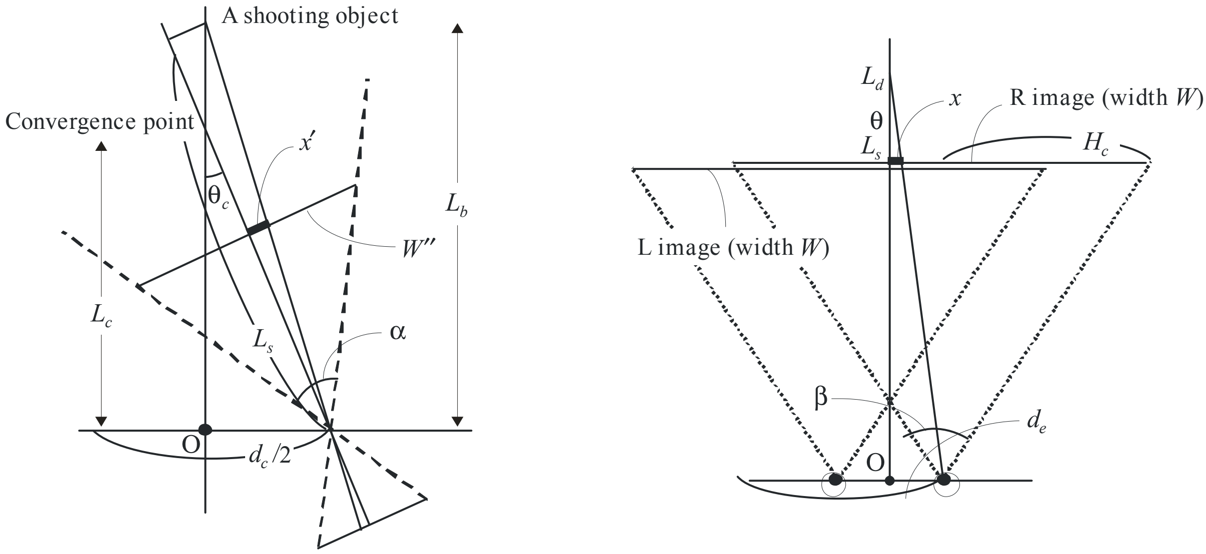
A basic requirement for the design of stereoscopic systems is an understanding of the transformation from real space (the space in which an actual object exists) to reproduced stereoscopic image space (the representation of this space in a stereoscopic image). In this section, we analyse the distortion of reproduced stereoscopic image space on the basis of image shooting and display system parameters[5].

2.1.1 Model of shooting/display systems

The configurations of the image shooting and display systems analysed here conform to the parameters shown in Fig. 3. The details of these parameters are shown in Table 1.

FIGURE 3

Shooting/display model



(a) Shooting system (b) Display system

Report BT.2160-03

Shooting and display systems can typically be configured in two different ways depending on how the optical axes are arranged.

Parallel configurations (where the two cameras of the stereo camera are aligned parallel to each other) are characterized such that objects at infinity are displayed at infinity by maintaining a constant horizontal separation of H_c between the left and right images when they are displayed (see Fig. 4). As a special case, when the separation d_c between the cameras and the horizontal offset H_c between the left and right images are equal to the separation d_e between the viewer's pupils, and the lens angle θ_b is equal to the angle of view of the display screen θ_d, the real space is in theory

---

5 YAMANOUE, H., OKUI, M. and OKANO, F., "Geometrical Analysis of Puppet Theater and Cardboard Effects in Stereoscopic HDTV Images", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 16, No. 6, p. 744-752, 2006.

reflected without distortion in the reproduced stereoscopic image space[6]. However, it is not always possible to satisfy this condition in broadcasting where a wide variety of different subjects are liable to be viewed under widely varying conditions.

Another optical axis configuration is the so-called toed-in configuration wherein the optical axes of the two cameras intersect (see Fig. 5). This configuration is characterized such that an object situated at the intersection of the optical axes appears at the depth position of the screen on which its stereoscopic image is displayed. It is also relatively easy to present a sense of depth for objects in the space in front of and behind the object at the intersection of the optical axes. By virtue of these characteristics, this method appears to be used in most stereoscopic programmes.

TABLE 1

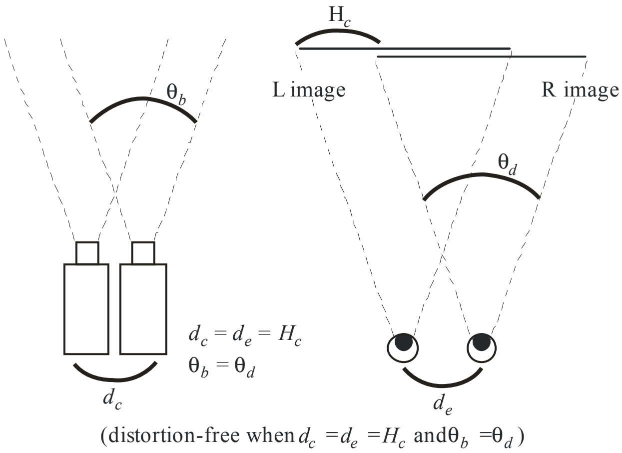
Parameters in shooting and display models

dc	Camera separation
de	Eye separation
Lb	Shooting distance
Lc	Convergence distance
Ls	Viewing distance
Ld	Position of a stereoscopic object
α	Angles of view of lens
β	Viewing angle
θc	Camera convergence angle
θ	Convergence angle of eye
Hc	Horizontal gap between L and R images
W	Width of screen
W'	Width of virtual screen at the viewing distance in the shooting model
x'	Distance from the centre of the virtual screen at the viewing distance in the shooting model (see Fig. 3)

[6] SPOTTISWOODE, R., SPOTTISWOODE, N. L. and SMITH, C., "Basic principles of the three-dimensional film", SMPTE J., Vol. 59, p. 249-286, 1952.

FIGURE 4

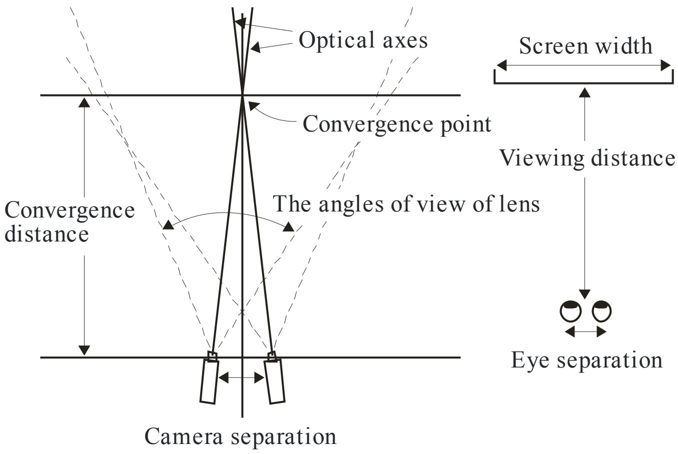
Parallel configuration



Report BT.2160-04

FIGURE 5

Toed-in configuration



Report BT.2160-05

2.1.2 Depth distance in real space and stereoscopic image space

If an object's depth position in real space (the environment where images are shot) and stereoscopic image space (the reproduced environment) are L_b and L_d, respectively, then the relationship between these values obeys the following formula using the parameters of Table 1 and the geometrical relationship of the system configuration shown in Fig. 3.

$$L_d = \frac{1}{\dfrac{1}{L_s} - \dfrac{a_1 \cdot a_2}{L_c} + \dfrac{a_1 \cdot a_2}{L_b} - \dfrac{H_c}{L_s \cdot d_e}} \tag{1}$$

In a parallel configuration, we can set $L_c \rightarrow \infty$ and $H_c = d_e$. In a toed-in configuration, we can set $H_c = 0$.

Table 2 shows the results of using equation (1) to investigate how the depth position L_d in the reproduced image and the actual depth position (the original camera-to-object distance) L_b are expressed in systems with parallel and toed-in configurations.

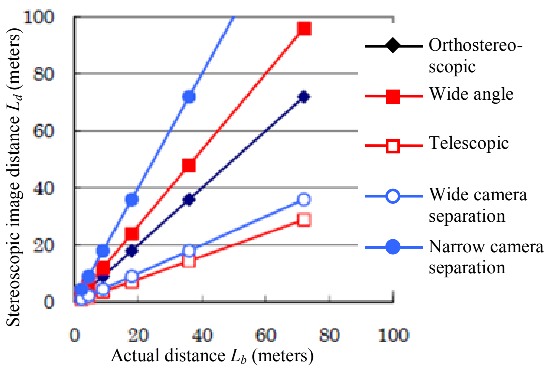
In Table 2, no consideration is given to the keystone distortion of the image shape that occurs in toed-in configurations. In other words, this table shows the characteristics at the centre of the image where keystone distortion has little effect.

In the parallel configuration, L_b and L_d obey a proportional relationship regardless of the parameter settings. On the other hand, in the toed-in configuration, L_b and L_d are equal only for a certain specific combination of parameters ($L_c = a_1 \cdot a_2 \cdot L_s$), but otherwise have a non-linear relationship. The graph of Table 2 indicates that different characteristics are exhibited depending on the sizes of L_c and $a_1 \cdot a_2 \cdot L_s$.

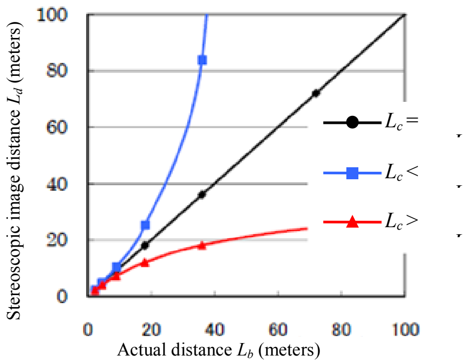
TABLE 2

Distances in real space and in reproduced stereoscopic image space

<table>
<tr><td>Parallel camera configuration
(Lens axes in parallel)</td><td>Toed-in camera configuration
(Converging lens axes)</td></tr>
<tr><td>$L_d = \frac{L_b}{a_1 \times a_2}$
(See the note below for a_1 and a_2)</td><td>$L_d = \frac{1}{\frac{1}{L_s} - \frac{a_1 \cdot a_2}{L_c} + \frac{a_1 \cdot a_2}{L_b}}$</td></tr>
<tr><td>



</td><td>



</td></tr>
<tr><td>Orthostereoscopic condition: When $\alpha=\beta$ and $d_c=d_e=H_c$, actual space is copied into a stereoscopic image by the same, or 100% magnification.</td><td>Nonlinearity is apparent when L_c is not equal to $a_1 \cdot a_2 \cdot L_s$.</td></tr>
<tr><td colspan="2">Note: $a_1 = \frac{d_c}{d_e}$, $a_2 = \frac{W}{W'} = \frac{\tan\frac{\beta}{2}}{\tan\frac{\alpha}{2}}$</td></tr>
</table>

2.2 Size distortion

The size of the reproduced image of an object shot by a camera at a certain distance (i.e. the size of recognized objects) varies with the image shooting conditions and display conditions, but is generally subject to size distortion. When size distortion causes objects to appear unnaturally small, it is often referred to as the "puppet theatre" effect[7].

---

[7] MACADAM, D. L., "Stereoscopic perceptions of size shape distance and directions", SMPTE J., Vol. 62, pp. 271-289, 1954.

2.2.1 Theoretical analysis

The puppet theatre effect has various interpretations. Here, we will not concern ourselves with the perceived absolute size of the reproduced image. Instead, we will concentrate on cases where the size of objects appears to be unnaturally deformed in comparison with the foreground and background.

If W_b is an object's size in the real space and W_r is its apparent size in the stereoscopic image space (i.e., its perceived size based on its depth position), then the relationship between these values is expressed by the formula below, where a_2 has the value shown in Table 2.

$$W_r = \frac{L_d}{L_b} \frac{\tan\frac{\beta}{2}}{\tan\frac{\alpha}{2}} W_b = \frac{L_d}{L_b} \cdot a_2 \cdot W_b \tag{2}$$

Now let us introduce the magnification M_s of the reproduced image. As shown in equation (3), M_s expresses the apparent size change of objects, taking their depth position into consideration. It corresponds to the ratio of the size W_r of the reproduced image to the size W_b of the object in real space.

$$M_S = \frac{W_r}{W_b} \tag{3}$$

Table 3 shows the results of applying M_s to the parameters of the parallel and toed-in camera configurations shown in Table 2. In the parallel configuration, M_s is constant regardless of the camera-to-object distance L_b, but in the toed-in configuration, M_s varies with L_b. In the toed-in configuration, the sizes of two objects at different depth distances (i.e. background and foreground objects) can be perceived very differently depending on the combination of parameters, and this is liable to cause puppet theatre effects in some cases.

TABLE 3

Analysis of magnitude distortion

Reproduction magnification ratio of an object (Ms)	
In the case of the parallel camera configuration: $Ms = \frac{1}{a_1}$	In the case of the toed-in camera configuration: $Ms = \frac{1}{\frac{1}{L_s} - \frac{a_1 \cdot a_2}{L_c} + \frac{a_1 \cdot a_2}{L_b}} \cdot \frac{a_2}{L_b}$
In the case of the parallel camera configuration, the reproduction magnification ratio is $1/a_1$.	In the case of the toed-in camera configuration, except where $L_c = a_1 \cdot a_2 \cdot L_s$, the reproduction magnification ratio of an object depends nonlinearly on the shooting distance and the background affects the perceived size of the object. Especially in the case of $L_c < a_1 \cdot a_2 \cdot L_s$, the Puppet Theater effect tends to be enhanced because the reproduced foreground is smaller than the background.

2.2.2 Subjective evaluation tests

In the discussion of the previous section, we showed that an object's reproduced magnification M_s changes with various different parameters. Thus, when two objects (e.g. foreground and background objects) are reproduced at different magnifications, it is expected that unnatural phenomena such as the puppet theatre effect will occur.

We tried to apply the geometrical discussion of the previous section to the results of subjective evaluation tests.

In the subjective evaluation tests, stereoscopic images were reproduced under a number of different imaging and display conditions, and the test subjects evaluated their subjective impressions of the object size[8].

For the evaluation images, we used images of a roughly life-sized mannequin (shown in Fig. 6) using three different configurations of shooting distances and camera lenses, each with three different distances separating the left and right cameras. This resulted in a total of nine different evaluation images. Table 4 lists the details of these imaging and display conditions. In these images, the foreground consists of the mannequin, and the background consists of a corridor doorway 4.5 m behind it. In each case, the depth position of the mannequin in the stereoscopic image was at the position of the screen, and the image was displayed life-size on a 120-inch screen.

The test subjects were asked to evaluate the subjective size of the reproduced stereoscopic image of the mannequin on the following five-grade scale:

5: Large

4: Somewhat large

3: Normal

2: Somewhat small

1: Small

FIGURE 6

Object used for subjective evaluation tests of puppet theatre effect

Report BT.2160-06

8 YAMANOUE, H., "The Relation between Size Distortion and Shooting Conditions for Stereoscopic Images", SMPTE Journal, Vol. 106, No. 4, pp. 225-232, 1997.

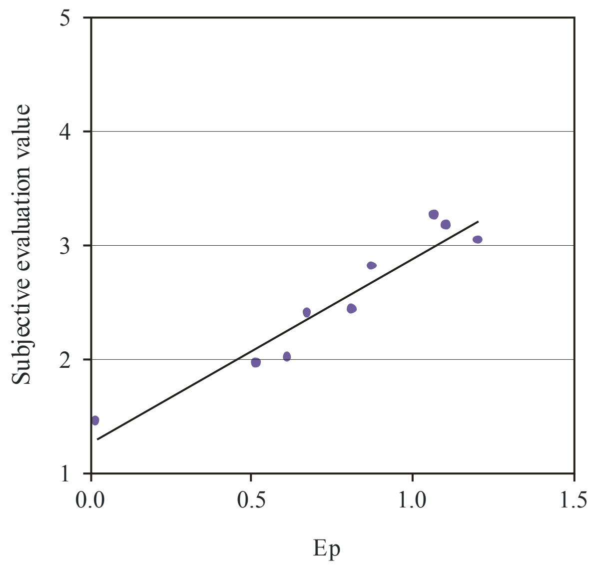
TABLE 4

Shooting and display conditions used in subjective evaluation tests of puppet theatre effect

conditions	1	2	3	4	5	6	7	8	9
contents	A			B			C		
the angles of view of lens α (degree)	51.3			27.0			18.2		
the distance to the foreground object (m)	3.0			6.0			9.0		
the distance to the background object (m)	7.5			10.5			13.5		
convergence distance Lc (m)	3.0			6.0			9.0		
camera separation dc (mm)	65	95	125	65	95	125	65	95	125
viewing angle β (degree)	33.4								
viewing distance Ls (m)	4.11								
a_1	1.00	1.46	1.92	1.00	1.46	1.92	1.00	1.46	1.92
a_2	0.63			1.25			1.88		
$Lc/(a_1*a_2*Ls)$	1.16	0.79	0.60	1.17	0.80	0.61	1.16	0.80	0.61
Ep	1.21	0.61	0.01	1.11	0.81	0.52	1.07	0.87	0.68

FIGURE 7

Correspondence with subjective evaluation scores



Report BT.2160-07

We used the ratio E_p of the foreground and background magnifications ($M_s^{(F)}$ and $M_s^{(B)}$) as a predicted value of the puppet theatre effect. Here, the depth position of the foreground and background are assumed to be given.

$$E_p = \frac{M_s^{(F)}}{M_s^{(B)}} \tag{4}$$

Figure 7 shows the correspondence between the predicted value E_p and the subjective evaluation scores. The evaluation scores and the reproduced image magnification ratio E_p show a strong correlation.

In practice, subjective size distortion phenomena such as the puppet theatre effect are thought to depend on the foreground and background and on the type of object depicted in the image, and it would be worth verifying these results with many more images. In this report, we have at least shown the geometrical criteria that cause this effect.

2.3 Depth distortion

Depending on the image shooting and display conditions, the relative size of the reproduced image in the depth direction may become smaller. In such cases, the apparent thickness of the reproduced object is distorted in the depth direction, causing a poor sense of depth. This phenomenon is sometimes called the cardboard effect[9] because it makes three-dimensional objects look like cardboard cut-outs.

2.3.1 Theoretical analysis

To quantitatively ascertain the thickness of a stereoscopic image based on the camera-to-object distance, we introduced the thickness term E_c expressed by the following formula:

$$E_c = \frac{\left(dL_d \middle/ dL_b \right)}{M_s} \tag{5}$$

Here, M_s is the magnification of the reproduced image, and L_b and L_d are the depth distances of the object in real space (where the object image is actually shot) and stereoscopic image display space, respectively. E_c expresses the degree of local changes in the depth direction.

Table 5 shows the results calculated by applying this value to the parameters of parallel and toed-in configurations shown in Table 2.

[9] HERMAN, S., "Principles of binocular 3D displays with application to television", SMPTE J., Vol. 80, pp. 539-544, 1971.

TABLE 5

Analysis of depth distortion

Reproduction magnification ratio of depth $E_c = \dfrac{\left(dL_d / dL_b \right)}{M_s}$	
In the parallel camera configuration	In the toe-in camera configuration
$E_c = \dfrac{1}{a_2}$	$E_c = \dfrac{1}{\dfrac{1}{L_s} - \dfrac{a_1 \cdot a_2}{L_c} + \dfrac{a_1 \cdot a_2}{L_b}} \cdot \dfrac{a_1}{L_b}$ Especially in the case of L_b=L_c (when it is reproduced on the screen) $E_c = a_1 \dfrac{L_s}{L_c}$

In the parallel configuration, E_c depends on a_2, which is simply the ratio of the angles of view when the image is shot and when the image is displayed. The situation is more complex in the toed-in configuration, where several factors are involved. However, when considering an object close to the depth position of the display screen, it can be seen that it is related to the camera separation (a_1), the distance to the intersection of the optical axes when capturing the image (L_c), and the viewing distance (L_s). It is suggested that these parameters should be considered in order to avoid the cardboard effect.

2.3.2 Subjective evaluation tests

The cardboard effect is a phenomenon whereby individual objects in a scene are perceived as having no depth, although it is generally still possible to ascertain the positional relationships among groups of objects in the depth direction. Although an analysis of the occurrence of this phenomenon is inevitably complex, here, we attempt to predict its occurrence in cases where the analysis is restricted to binocular parallax.

In the discussion of the previous section, we showed that the thickness E_c of an object's image is affected by various parameters. When the thickness E_c is small, it is predicted that unnatural phenomena such as the cardboard effect can occur.

We tried to apply the geometrical discussion of the previous section to the results of subjective evaluation tests. In the subjective evaluation tests, stereoscopic images were reproduced under a number of different imaging and display conditions, and the test subjects evaluated the subjective thickness of the object[10].

For the evaluated images, we produced images of the object shown in Fig. 8 under the nine sets of conditions shown in Table 6.

---

[10] YAMANOUE, H., OKUI, M. and YUYAMA, I., "A study on the relationship between shooting conditions and cardboard effect of stereoscopic images", IEEE Trans. Circuits and Systems for Video Technology, Vol. 10, No. 3, pp. 411-416, 2000.

The test subjects were asked to evaluate the subjective thickness of the reproduced stereoscopic image on the following five-grade scale:

5: Very thick

4: Thick

3: Somewhat thick

2: Not very thick

1: Not at all thick

FIGURE 8

Object used in the subjective evaluation tests of the cardboard effect

Report BT.2160-08

TABLE 6

Shooting and display conditions used in subjective evaluation tests of the cardboard effect

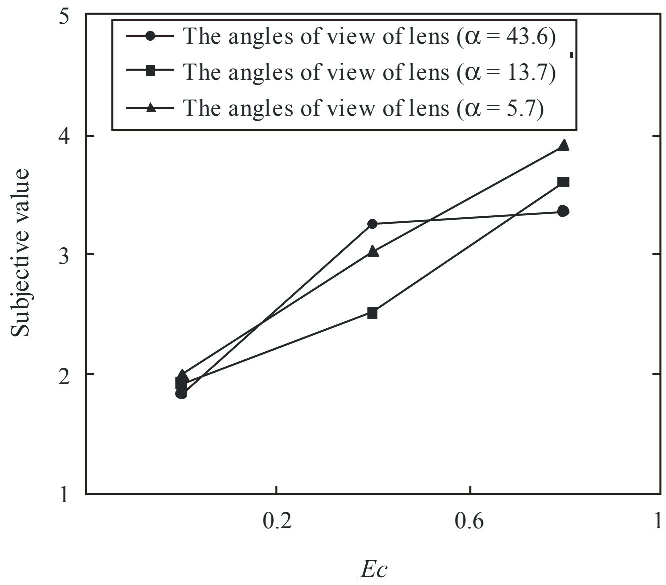
condition	1	2	3	4	5	6	7	8	9
camera configuration	toed -in								
the angles of view of lens α (degree)	43.6			13.7			5.7		
shooting distance Lb (m)	3.3			11.0			26.4		
convergence distance (m)	3.3			11.0			26.4		
camera separation dc (mm)	13	39	66	43	130	216	104	311	513
viewing angle β (degree)	43.6								
viewing distance Ls (m)	3.3								
a_1	0.20	0.60	1.02	0.66	2.00	3.32	1.60	4.78	7.89
Ec	0.20	0.60	1.02	0.20	0.60	1.00	0.20	0.60	0.99

Figure 9 shows the correspondence between the thickness E_c and the subjective evaluation scores. The vertical axis shows the results of normalizing the 5 evaluation categories in psychological space by using the method of successive categories. The calculation and subjective evaluations are clearly correlated.

The cardboard effect is a subjective effect that might be affected by other cues such as motion parallax and shading information. In this Report, we have at least shown the geometrical criteria that cause this effect.

FIGURE 9

Correspondence with subjective evaluation scores



Report BT.2160-09

3 Viewing comfort and discomfort of stereoscopic images

3.1 Parallax distribution and visual comfort of stereoscopic images

3.1.1 Introduction

One of the issues that arises when making stereoscopic images widely available to large numbers of viewers is the visual discomfort experienced by viewers of some scenes. This discomfort is thought to be caused by discrepancies and crosstalk between the characteristics of the images presented to the left and right eyes, such as differences in their geometrical characteristics or video properties, and is thought to be more of a problem in scenes where these characteristics are prominent. Another factor is parallax itself, which plays a key role in conveying the sense of depth.

It has often been noted that excessive parallax can cause visual fatigue. There are many productions where surprisingly large amounts of parallax are used intentionally for dramatic impact. If we can find out about the distribution of parallax within the same frame and how this distribution affects the viewer's visual discomfort, then this should provide us with very useful clues for the production of visually comfortable stereoscopic images in each scene. It can sometimes be difficult to analyse the parallax values included in a frame. In what follows, we describe such an analysis and apply it to a number of stereoscopic images. We then compare its results with those of subjective evaluation tests.

3.1.2 Parallax measurements

Here we summarize the parallax measurement method used in this study.

For many applications, we ideally need to be able to determine the depth (i.e. the amount of parallax) for each pixel of an image. However, algorithms for analysing parallax are generally prone to errors. Also, in practice there are cases where corresponding pixels do not exist in the images presented to the left and right eyes due to occlusion. In this study, we do not need to measure parallax on a strict per-pixel basis as long as we are able to extract the characteristics of a parallax distribution from the images. To achieve this aim, the method we propose combines phase correlation with a number of threshold processing methods. A detailed description of this algorithm can be found in the Reference below. The algorithm was used in the parallax analysis of stereoscopic images discussed below.

3.1.3 Subjective evaluation tests of parallax distributions and visual comfort

We performed subjective evaluation tests to investigate the relationship between visual comfort of stereoscopic images and their parallax distributions[11]. The subjective evaluation test conditions are shown in Table 7. The images used for the evaluation consisted of 48 different still images. These images were presented as stereoscopic images with 2D images as a standard reference, and their relative visual comfort was evaluated on a seven-grade scale. The 2D images were produced by presenting the left-eye image to both eyes and were evaluated by test subjects wearing the same polarizing glasses used for the stereoscopic images. The parallax in the stereoscopic images was measured by using the phase correlation method discussed in § 2. Here, the amount of parallax was measured on a per-pixel basis with the screen corresponding to a value of zero, and positions behind and in front of the screen correspond to positive and negative parallax values respectively. In the viewing conditions shown in Table 7, the amount of parallax of a single pixel corresponds to a separation of approximately 1 mm between the left and right images on the screen. Figure 10 shows the results of the visual comfort subjective evaluation tests and the results of measuring the amount of parallax in the images. In the graphs of this figure, the numbers on the horizontal axis designate the images to be evaluated. The graph at the top shows the results of the subjective evaluation tests. The vertical axis shows the mean value of the evaluation scores from 24 evaluators, and the vertical bars represent the standard deviations of these scores. The lower graph shows the results of measuring the amount of parallax. The vertical axis shows the amount of parallax measured in pixel units, the plotted points show the average values of the parallax in the images, and the vertical bars show the range of the parallax distributions. The upper and lower ends of the vertical bars represent the maximum and minimum parallax values. A comparison of the two graphs shows that the images with an evaluation score of 3 or less have a very large parallax distribution range.

TABLE 7

Conditions of subjective evaluation tests of parallax distributions

Images used in test	48 still images (including a standard pattern)
Subject	24 adult males and females (not expert)
Repeat test	10 s viewing of 2D image (for reference), following by 10 s viewing of stereoscopic image (for evaluation)
Display system	Stereoscopic HDTV using polarizing glasses

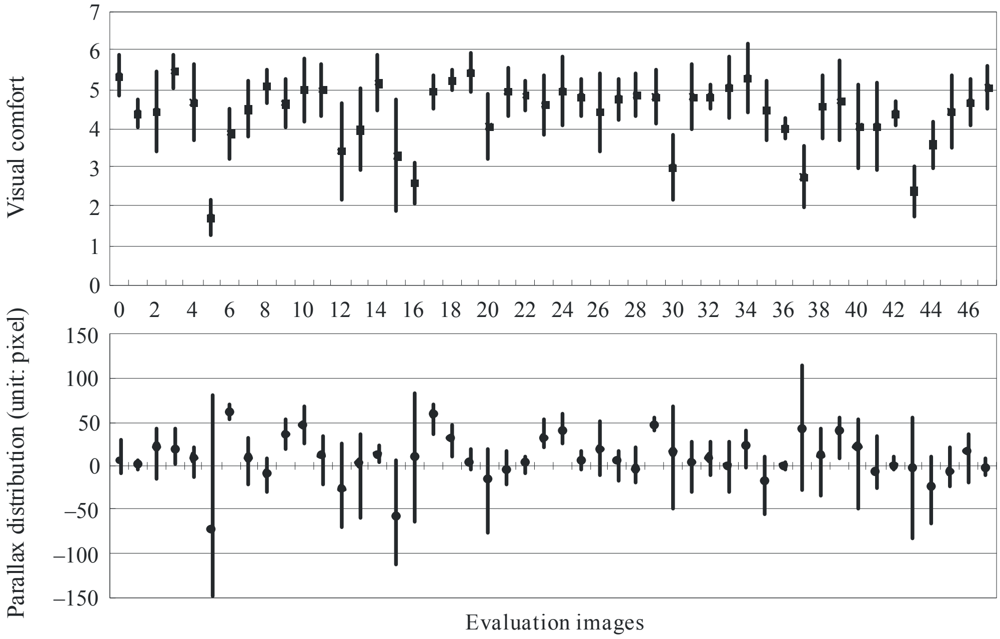
[11] NOJIRI, Y., YAMANOUE, H., HANAZATO, A. and OKANO, F., "Measurement of parallax distribution, and its application to the analysis of visual comfort for stereoscopic HDTV", Proc. SPIE, Vol. 5006, pp. 195-205, 2003.

TABLE 7 (*end*)

Screen size	90 inches
Viewing distance	About 3H (3.33 m)
Peak brightness	15 cd/m^2
Method of evaluation	Relative evaluation on a scale of seven, based on 2D image

FIGURE 10

Evaluation test results and parallax measurement results



Report BT.2160-10

On analysing the correlation between the results of the subjective evaluation tests relating to visual comfort and the statistical quantities (mean, range, minimum, maximum, variance) of the parallax distributions, it can be seen that the parallax range exhibits a strong correlation with a correlation coefficient of −0.86 (99% confidence).

Figure 11 shows the relationship between the parallax distributions and visual comfort of the images used in the evaluation. The vertical axis shows the amount of parallax in pixel units with a value of zero corresponding to the position of the screen, and the horizontal axis shows the visual comfort derived by using the method of successive categories to combine the psychometric values. The plotted points represent the mean values of the parallax distributions in the images, and the vertical bars represent the range of the parallax distributions. From Fig. 11, we can see that in each image evaluated as being visually comfortable, the range of the parallax distribution is approximately 60 pixels or less. This translates to a value of 0.3 diopters. Images are evaluated as being comfortable when the parallax distributions are in the range from approximately 30 pixels in front of the screen to approximately 65 pixels behind it.

Next, we investigated the relationship between the average values of the parallax distributions and the visual comfort of the images. With seven of the images, we performed visual comfort evaluation tests in which the average value of the parallax distribution was shifted to different positions. These tests were performed with 20 test subjects. The other test conditions were the same as in Table 7. Figure 12 shows the experimental results. In this figure, the points plotted with outlined symbols represent data obtained without horizontal shifting. As Fig. 12 shows, as the average value of the parallax distribution became closer to the screen position, the images were evaluated as being more visually comfortable.

FIGURE 11

Parallax distribution vs. visual comfort

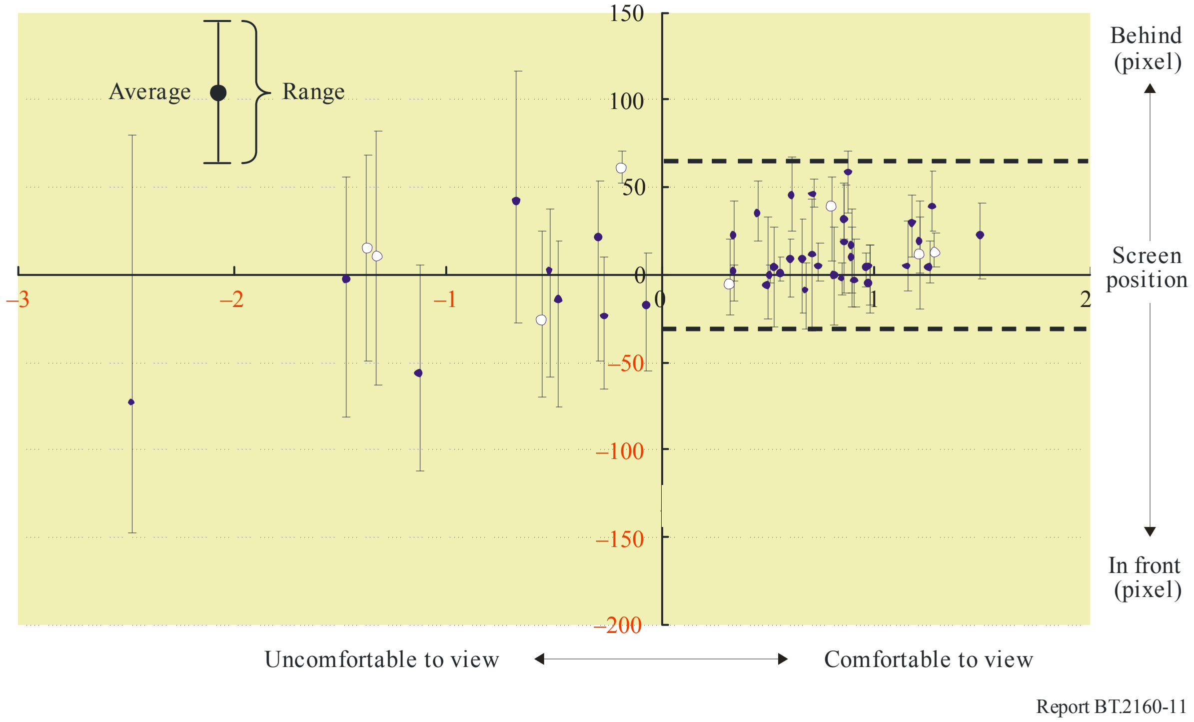
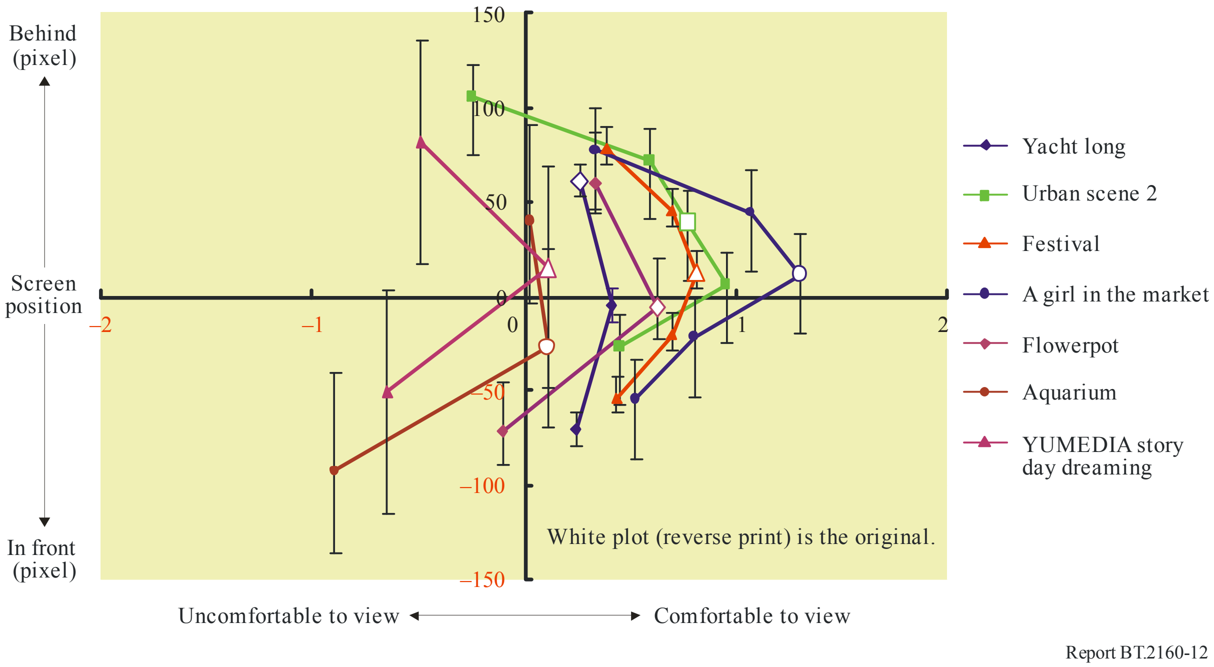


FIGURE 12

Average values of the parallax distributions vs. visual comfort



3.1.4 Subjective evaluation of the sense of presence

When scenes are limited to small values of parallax there might be a reduction in the positive effects of the stereoscopic images, such as the sense of presence. In the tests reported in § 3, the images were evaluated in terms of their sense of presence as well as their visual comfort. Specifically, the stereoscopic images were presented with 2D images as a standard reference, and their sense of presence was evaluated on a seven-grade scale.

In the analysis of the test results, we found no significant correlation between the sense of presence scores and the statistical quantities of the parallax distributions. We extracted the images for which the stereoscopic image was evaluated as more visually comfortable than the 2D image (35 images in total), and as a result of analysing these images, we showed that there is a strong correlation between the range of the parallax distribution and the sense of presence (correlation coefficient 0.65).

On the other hand, we observed no factorial effect of the average value of the parallax distribution on the sense of presence evaluation scores.

3.2 Visual comfort and discomfort in viewing stereoscopic images

3.2.1 Discrepancies between left and right images

Here we present the results of evaluating visual comfort/discomfort in viewing stereoscopic images from the perspective of discrepancies between left and right images in the performance or characteristics of image capture and display equipment. A series of experiments was conducted on discrepancies in size, verticality, inclination, and brightness, as well as on cross-talk regarding discomfort. Natural images of HDTV were mainly used for the evaluation. The major findings of the studies are summarized in Table 8. The values represent the results of a subjective evaluation test on the five-grade impairment scale.

When shooting and displaying stereoscopic 3DTV programmes, there can be geometrical distortions, such as size inconsistency, vertical shift, and rotation error, between left and right images [1]. Results of experiments where three kinds of distortions occurred independently from each other are shown in Table 8.

As differences between left and right images in amplitude and offset can be corrected but clipped white or black levels cannot, the degree of interference when the white or black level of one of the left and right images is clipped was evaluated. The evaluation results on detection and tolerance limits were reported to depend on image content [2]. Some of the earlier studies on differences in brightness and contrast are given by References [3] and [4].

Stereoscopic image cross-talk, in which the left and right images “leak” and can be partially seen by the opposite eye, can also result in discomfort for the viewer, and experiments evaluating it have also been done. These values on detection and tolerance limits were reported to be highly dependent on image content and display contrast, and cross talk must be reduced further still on high contrast displays [5]. One of earlier studies on cross-talk is given by Reference [6].

3.2.2 Depth range, distribution and change in parallax

Cases of extreme parallax or sudden changes in parallax cause visual discomfort, so it is important to manage parallax with special care when producing programmes with stereoscopic images. Results of experiments subjectively evaluating the relationship between visual comfort and the distribution of parallax in images, and the change in parallax before and after scene cuts are shown in Table 9. The positioning of subtitles has also been evaluated. All tests were done under standard viewing conditions for HDTV. The range of parallax for visual comfort has already been reported in § 3.1 of Annex 4.

TABLE 8

Research on impairment caused by discrepancies and cross-talk between L/R stereo images

<table>
<tr><th>Factor</th><th colspan="2">Image characteristic</th><th>Detection limit</th><th>Tolerance limit</th><th>Note</th><th>Reference</th></tr>
<tr><td rowspan="5">L/R image discrepancy</td><td rowspan="3">Geometric discrepancy</td><td>Size</td><td>1.2 %</td><td>2.9%</td><td>Taking size of one image as 100%</td><td rowspan="3">[1]</td></tr>
<tr><td>Vertical displacement</td><td>0.7%</td><td>1.5%</td><td>Taking image height as 100%</td></tr>
<tr><td>Rotation</td><td>0.5 deg.</td><td>1.1 deg.</td><td>Angle of rotation about image centre</td></tr>
<tr><td rowspan="2">Brightness discrepancy</td><td>White clip level</td><td>70%</td><td>60%</td><td>Taking 100 IRE white level as 100%</td><td rowspan="2">[2]</td></tr>
<tr><td>Black clip level</td><td>1%</td><td>2%</td><td>Taking 100 IRE white level as 100%</td></tr>
<tr><td>Cross-talk</td><td colspan="2">Contrast ratio of 100:1 in tests</td><td>1%-2%</td><td>5%-10%</td><td>Luminance ratio</td><td>[5]</td></tr>
</table>

TABLE 9

Research on parallax distribution, parallax range and parallax change over time

Factor	Evaluation item	Experimental results	Reference
Parallax distribution within image (frame)	Distribution pattern	Easily viewable if parallax is shaped far at top of image, near at bottom	[7]
	Range of parallax	Comfortable to view if parallax range is within 60 min.	[8]
	Subtitle position	Preferable if subtitle positions 10-15 min. in front of background image	[9]
Parallax change over time	Scene cuts	Uncomfortable to view if temporal change of parallax exceeds 60 min.	[9]

References

[1] H. YAMANOUE, M. NAGAYAMA, M. BITOU, J. TANADA, T. MOTOKI, T. MITSUHASHI, and M. HATORI, “Tolerance for Geometrical Distortions between L/R images in 3D-HDTV,” Systems and Computers in Japan, Vo1. 29, No. 5, 1998.

[2] A. HANAZATO, M. OKUI, and F. OKANO, “Evaluation of Stereoscopic Images Impaired by the Differences in Black/White Levels,” ITE Annual Convention 2001, 4-5, pp. 49-50 (2001) (in Japanese).

[3] J. FOURNIER and T. ALPERT, "Human factor requirements for a stereoscopic television service – Admissible contrast differences between the two channels of a stereoscopic camera," Proc. SPIE, Vol. 2177, pp. 45-54, 1994.

[4] P. BELDIE and B. KOST, "Luminance Asymmetry in Stereo TV Images, " SPIE, Vol. 1457, pp. 242-247, 1991.

[5] A. HANAZATO, M. OKUI, and I. YUYAMA, "Subjective Evaluation of Cross Talk Disturbance in Stereoscopic Displays," The 20th International Display Research Conference, pp. 288-291, Sep. 2000.

[6] S. PATOOR, "Human Factors of 3D imaging: Results of Recent Research at Heinrich-Hertz-Institut Berlin," Proc. of the 2nd International Display Workshops (IDW'95), Vol. 3, 3D-7, pp. 69-72, 1995.

[7] S. IDE, H. YAMANOUE, M. OKUI, F. OKANO, M. BITOU, and N. TERASHIMA, "Parallax distribution for ease of viewing in stereoscopic HDTV," Proc. SPIE, Vol. 4660, pp. 38-45, 2002.

[8] Y. NOJIRI, H. YAMANOUE, A. HANAZATO, and F. OKANO, "Measurement of parallax distribution, and its application to the analysis of visual comfort for stereoscopic HDTV," Proc. SPIE, Vol. 5006, pp. 195-205, 2003.

[9] Y. NOJIRI, H. YAMANOUE, A. HANAZATO, M. EMOTO, and F. OKANO, "Visual comfort/discomfort and visual fatigue caused by stereoscopic HDTV viewing," Proc. SPIE, Vol. 5291, pp. 303-313, 2004.

4 Visual fatigue in viewing stereoscopic images

4.1 Experimental results on inconsistency between vergence and accommodation

It has been confirmed through subjective evaluation that viewing stereoscopic images can result in a great degree of visual fatigue compared to viewing 2D ones. On this issue, it has been shown that changes in visual performance can be observed before and after viewing stereoscopic images, and that the fusional amplitude (the parallax range over which viewers can fuse left and right images) in particular decreases. The experimental result suggested the possibility that a fusional amplitude can be one of indices for evaluating visual fatigue [1].

Here, conditions in which vergence and accommodation are not consistent were reproduced using specialized equipment, and an objective evaluation of visual fatigue was obtained by measuring fusional amplitude before and after using these glasses for one hour. Results of these trials are shown in Figs 13 and 14. In both experiments, a common one-hour HDTV programme was used. The figures indicate mean values of ratios of the relative range of convergence (a relative value using the fusional amplitude before viewing as the basis), with small values indicating a narrower fusional amplitude.

The results from viewing a flat image in 3D for one hour with a fixed amount of parallax (such that the image is displayed in front or behind the screen) are shown in Fig. 13. The results of changing the parallax over time with the vergence and accommodation being consistent in one case, and not consistent in the other are shown in Fig. 14. Parallax was changed 16 times over a period of two minutes, and this was repeated 29 times. In both experiments, the same one-hour HDTV programme was used.

A large change in ratio value was observed when viewing images with parallax than when viewing 2D ones, and even more fatigue was caused by the time-varying images with inconsistencies between vergence and accommodation. These results indicate that inconsistencies between vergence and accommodation as well as fluctuations in time due to parallax, can be factors causing visual fatigue [2].

4.2 Experimental results on parallax amount and lateral/depth motion

A series of experiments have also been done to examine the relationship between this issue of inconsistency between vergence and accommodation, and depth-of-focus of the eyes.

First, the accommodation responses were measured [3]. Subjects viewed image content for approximately one hour on a 120-inch screen at a viewing distance of 4.5 m. The 3D image contents were two motion video sequences, and parallax for the video displayed under the conditions of the experiment was within the depth of field in almost all cases, 2D video was also used for reference.

The result was assumed significant for visual fatigue when the change of amplitude of the accommodation response in the before and after viewing was bigger than 0.5 diopters. From this aspect the results of comparing accommodation response before and after viewing showed no significant difference for the 2D images, whereas for the 3D images, three of five subjects for the first image sequence and two of five subjects for the second sequence showed visual fatigue. On the basis of these results and the fact that the video used was within the depth of focus, it is presumed that causes of visual fatigue other than depth of focus must also be considered.

Next, experiments were done to compare images in different amount of parallax, as well as 3D images with and without motion [4]. In the experiments, Japanese text with added parallax was displayed on a field-sequential 3D HD monitor (28-inch diagonal). From a viewing distance of 3H, subjects read for approximately one hour while turning pages using a mouse. This was repeated several times, changing the amount of parallax. Experiments were also done with moving text. Two types of text motion were tested: forward and backward in the depth direction, and horizontal motion. Here, the amount of parallax was limited within the depth of focus. The timing of motion was generated from an existing 3D programme.

The degree of fatigue was estimated on the basis of subjective evaluation and accommodation response. The subjective evaluation results are presented in Fig. 15. The figure shows there was a significant difference at the parallax of −1.36 degrees whereas at +1.36 degrees there was not, i.e. there was a large variance. Fatigue was also inferred when there was motion in the depth direction. These results suggest that changes in the depth direction can cause visual fatigue even when the amount of parallax remains within the depth of focus.

FIGURE 13

Fusional amplitude after viewing image with large binocular parallax

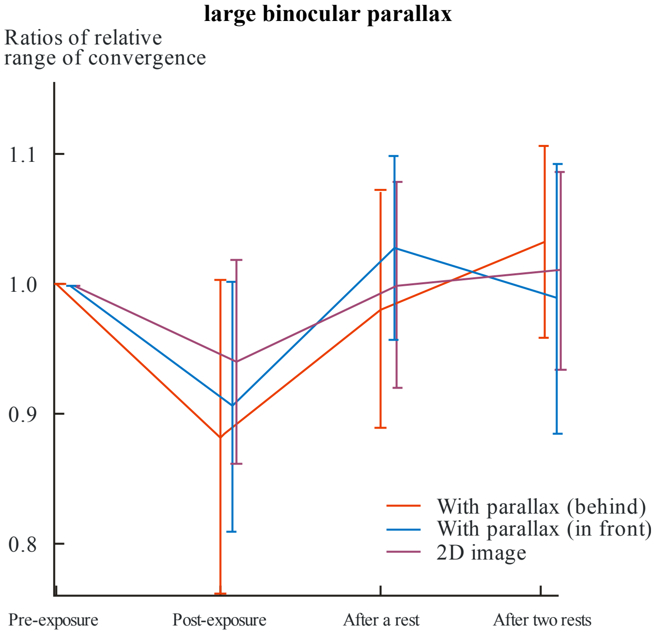


Figure shows mean value with standard deviation.

Report BT.2160-13

FIGURE 14

Fusional amplitude after viewing image with time fluctuations in binocular parallax

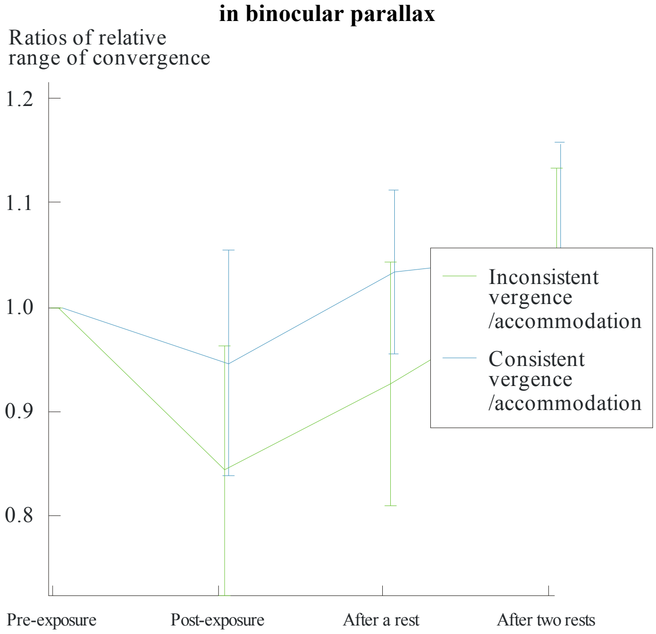


Figure shows mean value with standard deviation.

Report BT.2160-14

FIGURE 15

Subjective evaluation of visual fatigue while changing amount of parallax and stationary vs. moving objects

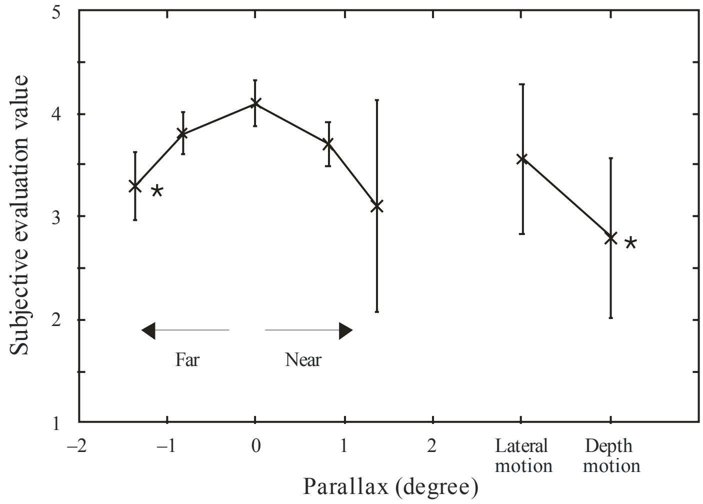


Figure shows mean values with standard deviation.
Asterisk indicates significant difference from still image on screen.

Report BT.2160-15

4.3 Evaluation of fatigue caused by watching 3DTV

4.3.1 Experiment

From January to March 2011, evaluation experiments were conducted consisting of 500 adult participants watching 3D content for approximately one hour on commercially available 46 to

50-inch 3DTVs that require the use of shutter glasses. The degree of fatigue after watching the 3DTV was evaluated under various viewing conditions.

The 3D content used in the experiment consisted of seven kinds of programmes (documentary, sports, music clip, animation, etc.) whose binocular disparities were mostly less than one degree. This content was recorded with a hard disk recorder (1 920 × 1080 60 i/10 bit/4:2:2) as 3DTV format referred to as Side-by-Side, where the horizontal resolution of the HD image was reduced by half.

Six types of viewing conditions were set, i.e., four viewing conditions in which participants watched 3D content with glasses in front of a 3DTV at three different distances (two, three and five times the screen's height) and from an oblique position of 40 degrees, and two other control conditions in which participants watched 2D content with or without glasses in front of a 3DTV at a distance of three times the screens height.

The participants were 500 women and men aged between 20 and 69 years old. Each participant watched 3DTV in one viewing condition only (Between-Group Design).

Visual acuity, Critical Flicker Frequency (CFF) [5], and the Advanced Trail Making Test (ATMT) [6] were used as objective indexes of fatigue, whereas the Simulator Sickness Questionnaire (SSQ) [7] and Visual Analogue Scale (VAS) [8] were used as subjective indexes. Fatigue caused by watching 3DTV was evaluated by the differences between those indexes evaluated before and after watching 3DTV. In addition, the participants answered a questionnaire about their physical conditions after watching TV programmes on the day when the experiment was conducted and the following day.

4.3.2 Results

The average value of each index before and after watching 3DTV was obtained for each viewing condition, and statistical analyses were conducted for testing whether these values show significant differences between different viewing conditions.

The results of the objective indexes indicated that there was no difference between watching 3DTV and traditional TV (i.e., watching 2D content without glasses) in degree of decline of visual and cognitive functions due to fatigue.

On the other hand, the results of subjective indexes indicated that there were some differences between watching 3DTV and traditional TV in the sensation of fatigue. However, these differences may not be attributed to watching 3D content, but to wearing the 3D shutter glasses.

Although there was no difference in the sensation of fatigue between the different conditions when evaluated immediately after watching 3DTV, the results suggest that the sensation of fatigue may be persistent if 3DTV is watched at a distance closer than the standard viewing position (i.e., three times the screen's height).

It should be noted that the results of the present study were obtained under conditions close to typical viewing situations at home, where the subjects simply watched 3D programmes whose binocular disparities were relatively small on commercially available 46 to 50-inch 3DTVs, and therefore these findings may not be applied to other viewing conditions and 3D content. Yano *et al*. [4], for instance, evaluated visual fatigue when subjects read the text of Japanese literature on a 28-inch CRT and reported that changes in depth direction can cause visual fatigue even when the amount of parallax remains within the depth of focus. Emoto *et al*. [2] evaluated visual fatigue when subjects counted the number of characters of Japanese translations superimposed on German opera and reported that there were differences in the P100 latency [9] of the visual evoked cortical potentials (VECP) between viewing 2D and 3D content, while the subjective evaluation revealed no difference.

References

[1] M. EMOTO, Y. NOJIRI, and F. OKANO [2004] Changes in fusional vergence limit and its hysteresis after viewing stereoscopic TV, Displays, Vol. 25, pp. 67-76.

[2] M. EMOTO, T. NIIDA, and F OKANO [December 2005] Repeated Vergence Adaptation Causes the Decline of Visual Functions in Watching Stereoscopic Television, Journal of Display Technology, Vol. 1, No. 2, pp. 328-340.

[3] S. YANO, S. IDE, T. MITSUHASHI, and H. THWAITES [2002] A study of visual fatigue and visual comfort for 3D HDTV/HDTV images, Displays, Vol. 23, pp. 191-201.

[4] S. YANO, M. EOMOTO, and T. MITSUHASHI [2004] Two factors in visual fatigue caused by stereoscopic HDTV images, Displays, Vol. 25, pp. 141-150.

[5] Critical Flicker Frequency is a psychophysical measure of visual temporal resolution. It represents the minimal number of flashes per second at which an intermittent light stimulus provides a continuous sensation, which may show some subject's visual fatigue and visual sensory sensitivity. Details of the method are described in the following paper:

T. MITSUHASHI [1996] Evaluation of Stereoscopic Picture Quality with CFF, Ergonomics, Vol. 39, No. 11.

[6] The Advanced Trail Making Test is a method to evaluate mental fatigue. In this test, circles numbered from 1 to 25 are placed randomly on the display and participants are required to use a computer mouse to click these circles in sequence. See the following reference:

O. KAJIMOTO: Development of a Method of Evaluation of Fatigue and its Economic Impacts. In Fatigue Science for Human Health. Edited by: Y. WATANABE, B. EVENGARD, B.H. NATELSON, L.A. JASON, H. KURATSUNE, New York: Springer; 2008:33-46.

[7] The Simulator Sickness Questionnaire (SSQ) has been used in many studies to measure the level of visually-induced motion sickness. The SSQ contains 16 questionnaire items with a four-point scale. See the following reference:

R.S. KENNEDY *et al*. [1993] Simulation Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness, The International Journal of Aviation Psychology, Vol. 3, No. 3, pp. 203-220.

[8] The Visual Analogue Scale measures subjective symptoms of fatigue where participants indicate the degree of fatigue on a simple visual analogue scale. "Guideline of Clinical Evaluation of Anti–fatigue" (in Japanese) cited this method as one of the standard evaluation methods of fatigue.

[9] P100 latency is a positive component with approximately 100-ms peak latency of the visual evoked cortical potentials (VECP) that was suggested as a fatigue index in the following study: T. YAMAZAKI, K. KAMIJO and S. FUKUZUMI [1990] Quantitative Evaluation of Visual Fatigue Encountered in Viewing Stereoscopic 3D Displays: Near-Point Distance and Visual Evoked Potential Study", Proc. SID, Vol. 31, No. 3, pp. 245-247.

5 Spatial distortion prediction system for 3DTV

5.1 Introduction

Spatial distortion of reproduced stereoscopic images is determined by a combination of factors including programme production techniques, display devices, 3D glasses, viewing conditions, and viewer characteristics. It is highly desirable to predict beforehand the degree and type of spatial distortion of the reproduced stereoscopic images so that more natural and more comfortable stereoscopic images can be presented to viewers.

This document describes a spatial distortion prediction system for a 3DTV (see Ref. (1)). This system calculates the spatial distortion of a reproduced stereoscopic image and predicts the extent of the puppet-theatre and cardboard effects, excessive binocular parallax, and excessive parallax distribution on the basis of the shooting, display, and viewing conditions.

5.2 Spatial distortion in 3DTV

Conditions under which images are captured, displayed, and viewed can contribute to the introduction of spatial distortions, that is, the differences between the real and the reproduced 3D spaces. Some spatial distortions might cause unnatural effects, such as the puppet-theatre effect and the cardboard effect. The puppet-theatre effect is an undesirable miniaturization effect that makes people look like animated puppets; the cardboard effect is a stereoscopic distortion causing an unnatural depth perception, where objects appear flat as if the scene is divided into discrete depth planes. When some objects are close to the camera, the entire stereoscopic image might appear to pop out from the screen and excessive binocular parallax and excessive parallax distribution may occur. Excessive parallax might also arise for background objects, which is known to cause visual discomfort or to prevent binocular fusion.

5.3 Spatial distortion prediction system for 3DTV

5.3.1 Use cases

A system capable of predicting spatial distortion and excessive parallax would be of great benefit to the industry, helping to provide more natural and more comfortable stereoscopic images.

Because a stereoscopic image can only be viewed properly under a particular viewing condition for most stereoscopic displays, the system can be used to select appropriate shooting parameters for a particular “standard” display/viewing environment. For a programme directed at children, the system might also be used to tailor the shooting parameters to their small interpupillary distance.

When a director intends to emphasize the reproduced depth to make objects jump out from the screen to have an impact on the viewers, the director must manage the shooting conditions to avoid spatial distortion that causes the puppet-theatre effect and excessive parallax distribution. It is particularly difficult to produce the intended stereoscopic images for large displays when shooting on location. This is because a small stereoscopic display or even a 2D display is often used to monitor the stereoscopicity at a close distance or to merely measure the horizontal disparities, resulting in the director choosing the shooting conditions more by trial and error than careful selection. The system would make it possible for the director to control the stereoscopicity accurately and easily.

In a 3DTV broadcast, a broadcaster might have control of the shooting conditions but has little control over the display and viewing conditions. On the other hand, the opposite is true for the viewer. Even so, the system might help identify suitable viewing conditions to recommend to viewers.

5.3.2 System outline

The system calculates the spatial distortion of a reproduced stereoscopic image and predicts the extent of the puppet-theatre and cardboard effects, excessive binocular parallax, and excessive parallax distribution on the basis of the shooting, display, and viewing conditions listed in Table 10. The relationship between the space to be shot (real space) and the space of the reproduced stereoscopic image (reproduced space) is calculated geometrically in terms of their depth and size.

The shooting conditions and the right and left images can be obtained from a stereoscopic camera system.

TABLE 10

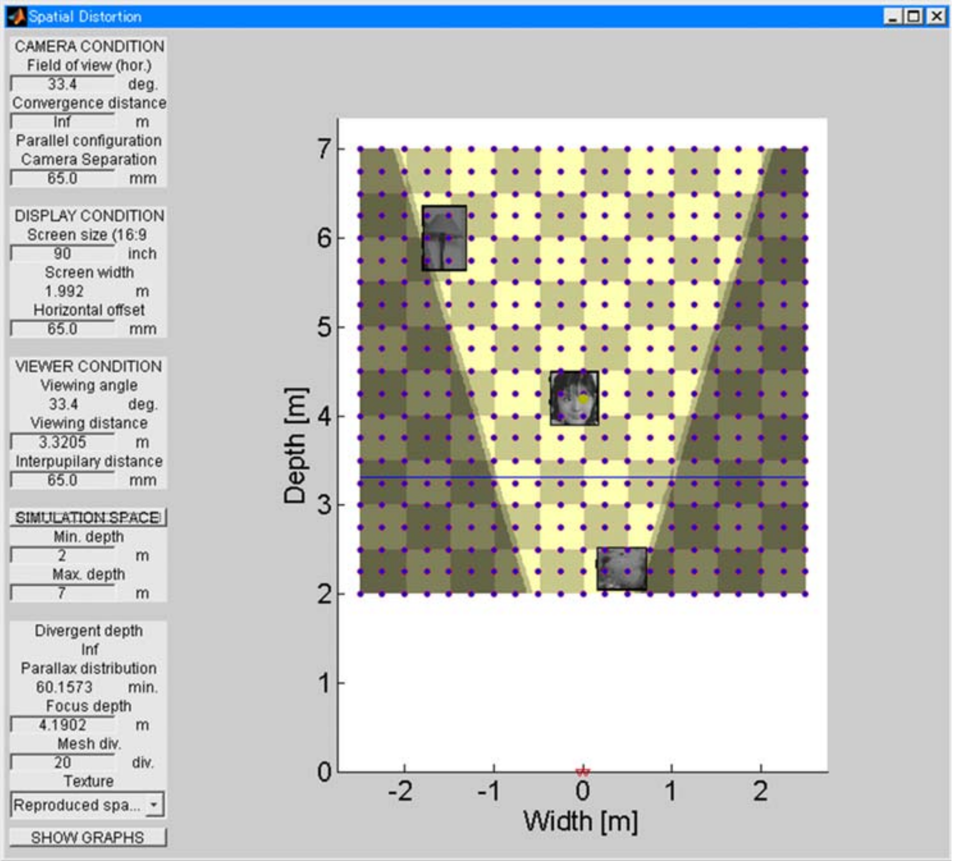
Parameters for shooting, display, and viewing conditions

Shooting parameters	Display parameters	Viewing parameters
Camera field of view Camera convergence distance Camera separation	Screen width Horizontal offset	Viewing distance Interpupillary distance

Figure 16 shows a ground plan of real and reproduced space grids without spatial distortion; the real space grid (shown by red dots) and the reproduced space grid (shown by blue dots and texture) coincide. It should be noted that the real space grid is always displayed as a square. The shooting, display, and viewing conditions are shown in the left pane of the window. The camera field of view and the display viewing angle are equalized, as are the camera separation, interpupillary distance, and horizontal offset. The system can measure the parallaxes of up to three objects, namely the object of interest, background, and foreground. Each object's depth in real space is calculated on the basis of the measured parallax and the shooting conditions. In order to determine the depth range of the space grid for calculating the spatial distortion, the user selects two portions, one at the maximum depth (a desk lamp in this example) and another at the minimum depth (a stuffed animal). A portion including the object of interest (a woman) should be selected to calculate the extent of the perceived puppet-theatre effect. The depth of the object of interest may be determined as the focus plane.

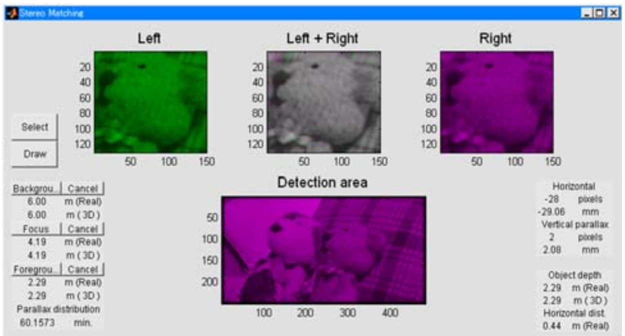
FIGURE 16

Screen shots of the spatial distortion prediction system



(a) Spatial distortion prediction window (under orthostereoscopic conditions)

(b) Captured image



(c) Object selection window

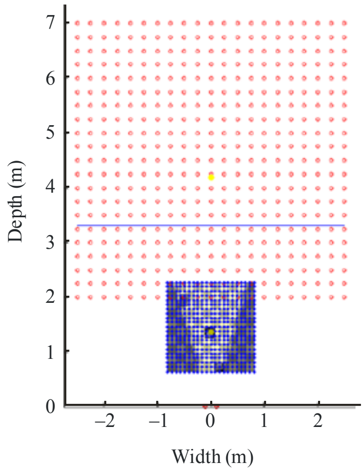
Report BT.2160-16

5.3.3 Examples of conditions and simulations

Figure 17 shows the simulation results obtained under four conditions for which some parameters were changed while other parameters remained the same as in Fig. 16. The extent of the perceived puppet-theatre and cardboard effects is expressed by changing the hue of the texture: magenta is increased in proportion to the extent of the perceived puppet-theatre effect and green is increased in proportion to the extent of the perceived cardboard effect. This system also produces an alert when excessive binocular parallax or parallax distribution is predicted.

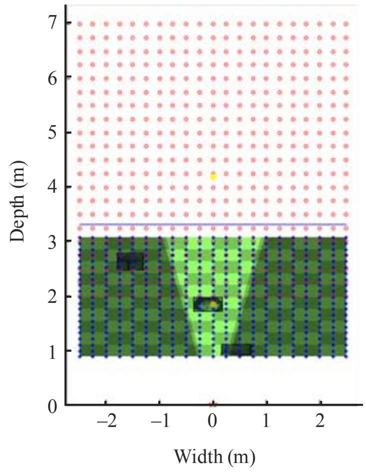
FIGURE 17

Simulation results



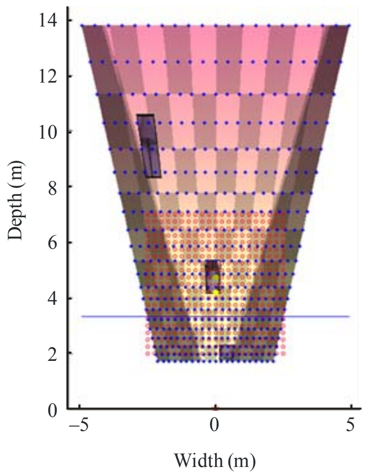
Case 1: Large camera separation of 200 mm

The reproduced space shrinks, and the objects appear to be small.



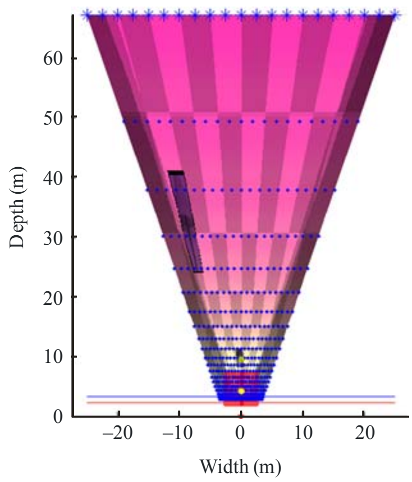
Case 2: Narrow camera field of view of 15º

The overall depth in the reproduced space is reduced, and a large cardboard effect is expected.



Case 3: Short interpupillary distance of 45 mm

A slight puppet-theatre effect is expected, and the object of interest (woman) appears small.



Case 4: Short convergence distance of 2.3 m

The puppet-theatre effect is expected. Objects at a depth of approximately 7.5 m are reproduced at infinity.

Report BT.2160-17

NOTE – Magenta colour is increased in proportion to the extent of the perceived puppet-theatre effect. Green colour is increased in proportion to the extent of the perceived cardboard effect.

References

(1) K. MASAOKA, A. HANAZATO, M. EMOTO, H. YAMANOUE, Y. NOJIRI, F. OKANO, "Spatial distortion prediction system for stereoscopic images," Journal of Electronic Imaging 15(01), 013002, 2006.

Annex 5

3DTV Broadcasting Safety Guideline in Korea

1 Necessity of Safety Guideline

1.1 Necessity

Unlike 2DTV, 3D TV might cause some sort of discomfort such as dizziness, headache or fatigue of eyes to some viewers. Thus it is necessary to prepare a 3DTV broadcasting safety guideline to minimize or remove the cause of such discomfort so that safer and more comfortable viewing of 3D broadcasting services is ensured.

1.2 Typical discomforts

Fatigue of eyes

- Unclear vision (eyes being dimmed)
- Eyes feeling heavy
- Double vision
- Eyes being dried

Fatigue of body

- Headache or migraine
- Sickness and nausea
- Dizziness

2 Viewing circumstances guideline

2.1 Viewing time and rest time

Viewers are recommended to take a 5~15 min break every hour, which is also recommended for 2D video display [1].

2.2 Viewing distance

The recommend viewing distance is 3~6 times of the height of 3D video display. For example, 2~4 m is recommended for 55 inch TV. If there is not sufficient space for the above distance, viewers are recommended to view TV at the farthest distance possible.

– In case of 2D video display, viewing at too short distance deteriorates the perceivable spatial resolution, and viewing at too far distance reduces absorption. For 3D video display, another consideration is the size of binocular disparity. Viewing an image closer raises the size of binocular disparity entered into the viewer’s eyes, causing optical discomfort [2].

2.3 Viewing position

2.3.1 Viewing position

Viewers are recommended to take the position to face the centre of the video right in front of the display.

– Viewing 3D video right in front of the display minimizes distortion of scenes. Viewing the video at an oblique angle causes “shear distortion” which is a phenomenon that 3D image is distorted as if it follows the viewer. It also distorts the shape and the size of image formed on the eyes. These distortions may cause optical discomfort.

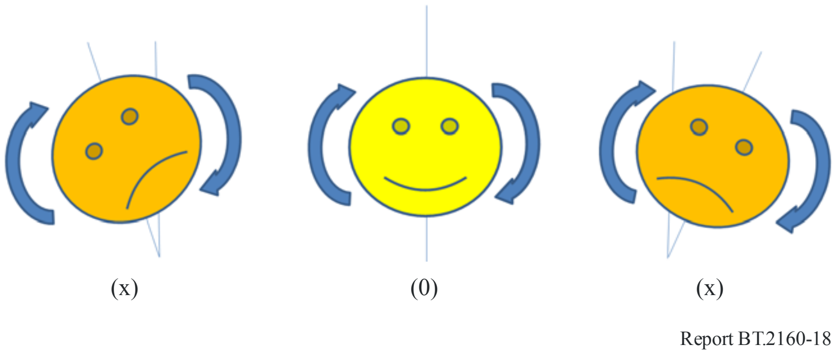
2.3.2 Horizontal viewing position

Viewers are recommended to maintain level of the display and eyes.

– Binocular disparity is given horizontally to the display. If the head of a viewer is inclined considerably to a side, the binocular disparity of the image is entered as a vertical parallax into the viewer’s stereopsis system. Therefore, the viewer feels difficulty in perceiving the depth provided by the binocular disparity. Even if the viewer manages to perceive the depth, the vertical parallax increased by the inclined head makes it difficult for the viewer to fuse two images into a single 3D image.

FIGURE 18

Horizontal viewing position



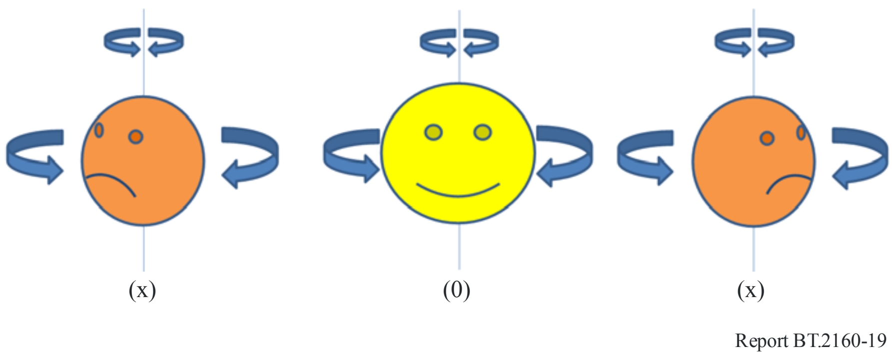
2.3.3 Right angle viewing position

Viewers should stare the display at the right angle in front of the display.

– Viewing a display with the head turned to a side makes a difference of size of image entering into eyes, causing a difficulty in fusing two images into a single 3D image.

FIGURE 19

Right angle viewing position



2.4 Others

Viewers are recommended to maintain comfortable viewing conditions by taking the following actions [1]:

– Adjust light, sound and air condition of the viewing space.
– Adjust brightness and colour of the display.

3 Viewer guideline

3.1 Symptoms caused by 3D viewing on viewers

Viewers are recommended to stop viewing 3DTV if they have a headache, pain in eyes, dizziness, nausea, palpitation, unclear vision, unpleasant feeling, optical discomfort or double vision [1].

Fast movement of objects on the display or excessive accommodation-convergence mismatch/conflict may cause fatigue of eyes. In this case, viewers should stop viewing 3DTV and take a break. If viewers experience any other discomfort, they should stop watching and stare into a far distance.

3.2 Stereo blindness and stereo abnormality

Viewers with stereo blindness or stereo abnormality may suffer an optical discomfort with double vision. They are recommended to avoid watching 3D images.

– Not every person can perceive vivid stereoscopic depth from 3D images. According to the report, about 1% of population cannot perceive the stereoscopic depth for any reason, and about 30% can perceive the stereoscopic depth by detecting the stimulation for a specific depth only from the protruded or retreated images [3][4][5].
– Viewers having a strabismus or astigmatism may have a difficulty in perceiving the stereoscopic depth, and suffer more severe fatigue of eyes with double image than the viewers with normal vision.
– Amblyopia is accompanied by astigmatism in many cases for children, and mostly, they cannot perceive the stereoscopic depth. Viewers with large difference between left and right eye may have a difficulty in deceiving the stereoscopic depth.

3.3 Chronic diseases

Viewers with any chronic disease (epilepsy, cardiac disorder, high/low blood pressure, etc.) are required to pay special attention when viewing 3D images.

3.4 Age

For children under 10 it is recommended to control the viewing conditions or to give guidance on viewing 3D images, since they are in the course of development of optic system and functions, [1].

– Children have relatively shorter distance between eyes, and they perceive higher binocular disparity of 3D image than adults. Therefore, children perceive higher stereoscopic depth than adults from the same 3D image.

Ageing may deteriorate the stereopsis function.

– It is reported that ageing reduces the optical capability of deceiving the stereoscopic depth. Therefore, middle-aged/old viewers may have a difficulty in perceiving a vivid stereoscopic sense when compared to the younger age [6].

4 Content guideline

4.1 Setting stereo cameras

When producing stereoscopic image with stereo cameras, the main element is to set the parameters consistently between left and right cameras. For this purpose, the following considerations are recommended:

– It is important to adjust the optic axes of the stereo cameras. The vertical adjustment error, rotation adjustment error and intersection error must be minimized.

– The basic camera parameters must be set so that there is no error of zoom, focus, iris and colour between right/left images. It is recommended to have left and right cameras synchronized.

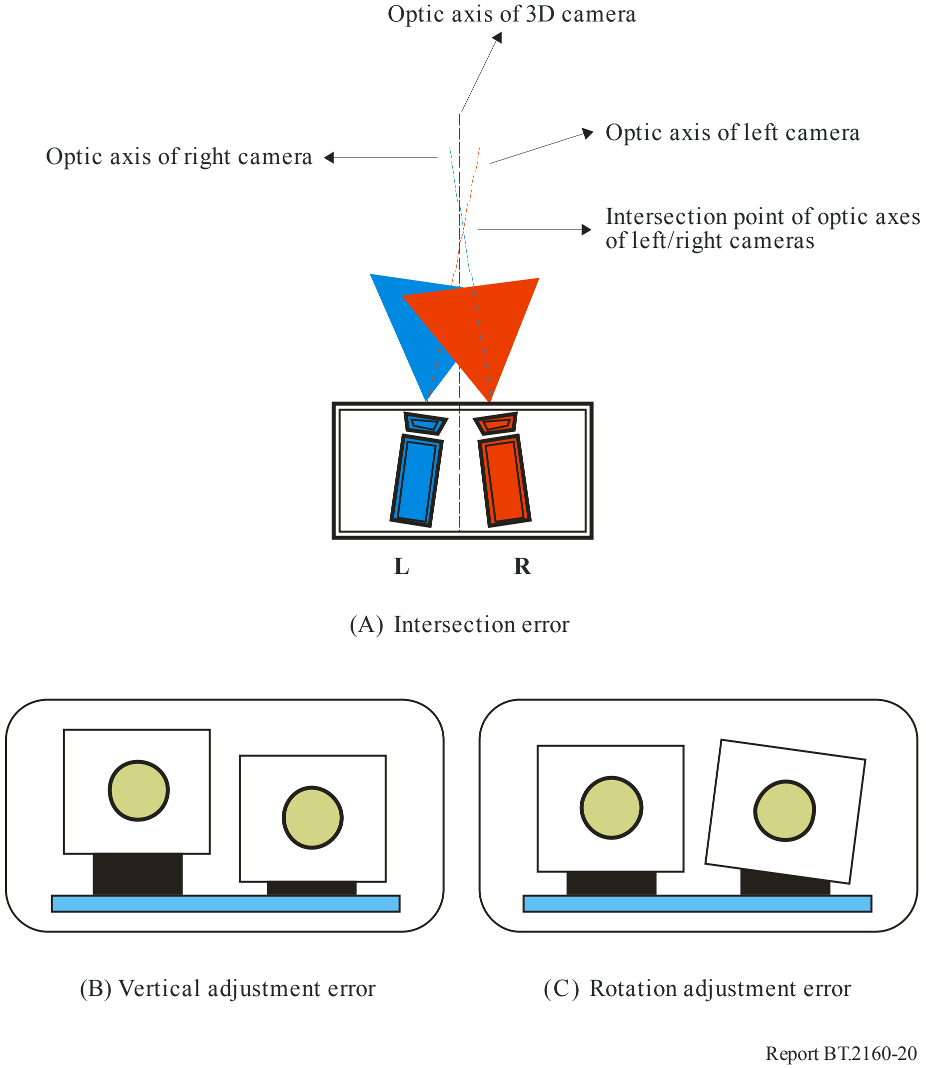
– Vertical inconsistency of images caused by inconsistency of optic axes is known to cause a fatigue of eyes.

– The state that the intersection point between optic axes of left and right cameras is not on the centre lines of the stereo cameras is called an intersection error. An intersection error may occur due to optic axes adjustment error even if the stereo cameras are set properly without any vertical adjustment error or rotation adjustment error. If the intersection error is excessive, the cameras cannot express the symmetric stereoscopic sense on the left/right symmetric stage, but express in appropriate stereoscopic sense.

– Excessive stereoscopic sense continued unnecessarily or sudden change of stereoscopic sense in the course of production of content may easily cause a fatigue of eyes.

FIGURE 20

Types of error of optic axis adjustment error



4.2 Taking stereoscopic images

Following considerations are recommended in taking 3D stereoscopic images.

- Images must be produced in the manner that fatigue of eyes caused by sudden change of vision should be minimized. For this purpose, images should be produced in the manner that no sudden change of disparity should occur in a short period of time, and it is recommended to prevent fatigue of eyes by adopting smooth camera work in using zoom, panning and other display techniques. It is also recommended that the images should be edited to prevent a severe change of disparity between scenes.
- 3D image distortion, which does not occur in 2D images, should be minimized. When taking images closely to an object with cross cameras, one must be careful not to make any keystone distortion. A compensating work is recommended if the eyes are not comfortable with the images.
- In taking images, one should be careful in using zoom lenses, and it is recommended to minimize the cardboard effect and the puppet theatre effect by adjusting the space between cameras.
- When presenting an object near a side edge of the display, the object should not be projected if possible.

- If the depth of an object perceived through binocular screen disparity is too small when compared to the known depth of the object, a cardboard effect may occur. In this case, this effect should be minimized with the camera work [8].
- A puppet theatre effect is made when there is a gap between the stereoscopic image displayed and the size of the object perceived in the real world. Viewers are likely to perceive an object as small as a doll. This effect easily occurs especially on a small display [8].
- A disparity between left and right images should be consistent if the image is produced properly. If a scene is taken with cross stereo cameras, the image produces a keystone distortion in the echelon formation with inappropriate vertical parallax. The vertical parallax increases as the space between cross stereo cameras grows and the focal distance of the lens reduces [9].
- Cross stereo cameras may create an inappropriate horizontal parallax, which in turn, causes a distortion of stereoscopic depth. Viewers may perceive that an object at the edge is farther than that in the centre of the display [9].

4.3 Caption

In 3DTV, the caption must be displayed in front of objects in order to prevent that the caption is unnaturally inside an object. If an object is projected, however, the caption being nearer to viewers than the object may cause a fatigue of eyes due to an excessive disparity. In this case, it is recommended to change the position of the caption to prevent fatigue of eyes.

4.4 Screen disparity

- For uncrossed disparity, the screen disparity between left/right images should not exceed the average gap between eyes.
- For crossed disparity, excessive screen disparity between left/right images may cause a fatigue of eyes.
- Continued viewing of excessive binocular disparity causes a fatigue of eyes. Therefore, it is recommended not to present excessive crossed disparity continuously.

5 Display guideline

5.1 Crosstalk of display

A crosstalk of display causes discomfort in viewing stereoscopic images. It is recommended to minimize crosstalk in consideration of the following:

– A crosstalk of display is a phenomenon occurring due to incomplete separation of left and right images. It varies depending on the display method and left/right image separation method. For example, in the polariscopic method, crosstalk is caused by the optical performance of the polarizing filter and the incomplete adjustment between the polarizing filter and the display pixel elements. In case of active shutter method, the pixel response speed becomes the main factor of crosstalk.

– Crosstalk of display is an independent parameter to the content, and can be indicated in an objective value for each display. For example, the crosstalk for the left eye is induced from the brightness of the right image versus that of the left image seen by the left eye, and vice versa.

The crosstalk experienced subjectively by viewers may be caused by the content, as well as the display crosstalk. Therefore, it is recommended to consider the following when producing a content:

– Subjective crosstalk experienced increases if there is a large contrast or binocular disparity between left and right images at the same position of the display [7].

5.2 Display refresh rate

In order to prevent flickering, the refresh rate of left and right images should be 60 Hz or higher. Therefore, the total refresh rate of a time-division 3D display should be 120 Hz or higher.

5.3 3D glasses

In addition to the display crosstalk, crosstalk may occur due to the glasses. Therefore, it is recommended to consider the following to minimize crosstalk:

– For polarized glasses, leakage of light caused by optical performance of the polarizing filter generates crosstalk. For active shutter glasses, crosstalk is generated by the light penetrated due to the optical performance of the glasses when the liquid crystal is closed, and the sync timing gap between the display and the glasses.

– Active shutter glasses acquire information on synchronization between left and right images and the action commands are acquired via communication with the display. In order to acquire an intensive image effect with 3D glasses at conversion from 2D to 3D and during viewing 3D images, the signals and protocols for communication shall be robust against external interference, and interruption shall be minimized.

References

[1] ISO IWA3, Image safety-Reducing the incidence of undesirable biomedical effects caused by visual image sequences, 2005.

[2] SUGAWARA, M., MITIANI, K., KANAZAWA, F.M., OKANO, F. & NISHIDA, Y., "Future prospects of HDTV – Technical trends toward 1080p", Paper presented IBC2005 conference, 2005.

[3] VAN EE, R., "Correlation between stereoanomaly and perceived depth when disparity and motion interact in binocular matching", Perception, 32, pp. 67-84, 2003.

[4] RICHARDS, W., "Stereopsis and stereoblindness. Experimental Brain Research", 10, pp. 380-388, 1970.

[5] RICHARDS, W., "Anomalous stereoscopic depth perception", Journal of the Optical Society of America, 61, pp. 410-419, 1971.

[6] BELL, B., WOLF, E., & BERNHOLZ, C.D., "Depth perception as a function of age", Aging and Human Development, 3, pp. 77-81, 1972.

[7] WOODS, A.J., "Understanding crosstalk in stereoscopic displays", Keynote presentation in 3DSA conference, Tokyo, Japan, pp. 19-21, 2010.

[8] Lydia M.J. MEETSTERS, Wijnand A. IJSSELSTEIJN, and Pieter J.H. SEUNTIENS, "A Survey of Perceptual Evaluations and Requirements of Three-Dimensional TV", IEEE Transactions on circuits and systems for video technology, vol. 14, no. 3, pp. 381-391, 2004.

[9] Andrew WOODS, "Image Distortions in Stereoscopic Video Systems", Stereoscopic Displays and Applications IV, Proceedings of the SPIE Vol. 1915, San Jose, CA, 1993.

Annex 6

Italian Health Ministry Circular Letters

The Italian Health Ministry issued a Circular Letter dated March 17, 2010, addressed to all the Regional Health Agencies, the Police, the associations of cinema operators and for information to the Ministry for Economy Development, which is responsible for telecommunications including broadcasting. The Italian Ministry action was prompted by a request of the Italian Consumer Protection Authority, and it was based on advice provided by the Ministry's High Advisory Council on Health. Some further clarifications were issued in a subsequent circular letter dated 6 August 2010.

The circulars are available (in Italian) on the website of the Ministry, and a translation is provided below.

The Ministry circular letter of March 17, 2010 states that:

– the scientific literature does not seem to provide firm proof that viewing stereoscopic programming would force human eyes and brain to perform unnatural processing of visual information; consequently there are no clinical indications at this moment against the use of 3D spectacles during cinema screenings, on condition that such screenings are limited in duration (the advice of the High Advisory Council on Health on this point was more detailed: it is suggested that viewing should be limited in time, and that it should contain intermissions proportionate to the total programme duration);

– some functional problems may arise in young viewers due to the use of 3D spectacles to view cinema presentations, because binocular vision may not yet be present or well established in young viewers, or because they may be cross-eyed or amblyopic, or because they may be going through a period of *visus* rehabilitation; however those problems should cause no irreversible damage or pathologies;

– consequently, the public that attends stereoscopic screenings should be informed that children under six years should not use 3D spectacles and that even adults should not use them for a duration that exceeds a single screening session.

A further circular letter of the Ministry, dated 6 August 2010, stated that in those cases when single-use glasses cannot be considered due to their technology or cost, 3D glasses must be properly disinfected and repackaged before they are provided to the next user since the risk of transmission of bacterial infections tends to increase with the successive use of the same spectacles by different viewers.

Annex 7

Example notifications given to viewers in Japan

Background

The following notifications prepared by DPA[12] are presented to viewers on the air before stereoscopic 3DTV programmes are broadcast or by the web site.

1 Notifications when 3D programmes are broadcast on same channel as 2D programmes

- This programme is 3DTV. (When a 3DTV programme starts.)
- The 3DTV programme is about to end, and will be followed by a 2DTV programme. (When a 3DTV programme ends.)
- Change the “3D/2D” mode as appropriate. (When the type of programme, i.e. 3D or 2D, changes.)

2 Notifications that should be broadcast with 3DTV programmes

- Watch TV in a well-lighted room and at an adequate viewing distance.
- Stop watching, take off your 3D glasses, and take a rest if feeling discomfort while watching 3DTV.
- Supervise infants’ viewing of 3DTV.
- Check whether stereoscopic images can be correctly seen or not by using inserted clips before starting 3DTV programmes.

3 Notifications may inform viewers, even though some of these are basically in the product manual

- Prepare appropriate products for 3DTV to use in programmes.
- Side-by-side images may be shown on the display when watching 3DTV programmes on 2D television.
- The On Screen Display (OSD) may disrupt 3D images.
- Refer to safety precautions in using 3D eyewear.
- Information on content of 3DTV programmes is available from the electronic programme guide (EPG) or homepage to help with programme choices. `

---

[12] DPA (The Association for the Promotion of Digital Broadcasting) is an organization of Japanese broadcasters and manufacturers associated with digital TV.

Annex 8

Technical Report 10

3D Briefing Document for Senior Broadcast Management

EBU TECHNOLOGY AND DEVELOPMENT

Your reference in media technology and innovation

3D TV - Its importance to EBU Members

Geneva
December 2010

EBU·UER

Contents

[*] Issued by EBU TECHNOLOGY & DEVELOPMENT
L'Ancienne Route 17A
1218 Le Grand-Saconnex
Switzerland

tech@ebu.ch

3D TV Briefing Document
for Senior Broadcast Management

Keywords: 3D television, Stereoscopy, Displays, Glasses, 3D Production.

3D TV Synopsis – What is important for EBU Members?

The value of 3D to the audience

The perception of the public towards 3D is predominately determined by cinema experiences. In spite of this, the consumer electronic manufacturers are promoting 3D as the next big desirable viewing experience in the home.

The EBU believes that public service broadcasters (PSBs) must take a pragmatic approach to 3D services and they should be aware of the value of an event or programme to its audience when making decisions about producing and broadcasting in 3D.

Programmes such as the Olympics and Eurovision Song Contests can be regarded as an "appointment to view" where an audience will be prepared to use glasses and watch 3D content instead of a 2D simulcast. This may not be the case with day to day viewing where the 2D version will just seem easier to watch.

Certain content genres such as wildlife documentary, especially macro and animation elements and some live events in smaller size areas can deliver a very immersive 3D viewing experience.

The display market

It is expected that all larger displays will be 3D capable in the future. For PSBs it is important to understand the potential number of households that are equipped with 3D capable displays.

GfK figures (Oct. 2010) suggest:

It is expected that by 2014 about 42% of all purchased display devices will be 3D capable.

The number of 3D devices in Europe is expected to be 600 000 by the end of 2010, and by the end of 2011, about 3 Million devices.

The primary question is whether consumers will accept 3D at all and whether they will invest in additional glasses (usually the displays are sold with one or two pairs of 3D glasses).

Content

Creating good 3D content requires special equipment and skills.

The current position of 3D rights is unclear, and public service broadcasters must pay close attention to developments in the discussions of 3D rights.

PSBs should grasp their responsibility to inform and educate the public about 3D for Television (in a way similar to that done for HDTV).

3D content is best produced with a particular range of screen sizes - home (30" to 70"), cinema or IMAX in mind.

Stereographers use different depth budgets for cinema and TV. Disparity (see glossary) that produces a good but large 3D effect on TV could cause extreme 'pain' as your eyes try to look in opposite directions in the cinema!

Making 3D that is OK on all screen sizes may be just that; OK but not stunning. This means that 3D content produced for the cinema requires post-processing to work adequately on consumer sized displays (and vice-versa). Not following this guideline will generally lead to unsatisfactory results.

3D and Health Issues

Poor stereoscopy *is* responsible for headaches, eye strain and nausea!

Estimates[1] of the number of people who suffer from 'stereo blindness' vary from 5% to as many as 15% of the population!

> *Note:* *Stereoscopic 3D-TV displays call on the eye-brain to work in a way they do not normally do by separating the functions of focussing and pointing. This can cause eye discomfort, and may cause eye fatigue or other symptoms. The ITU-R is considering these 'health-related' issues. There is generally a shortage of information about the potential short or longer term effects of viewing S3D. The ITU-R, prompted by the views of the Italian administration, also recognizes that for environments such as the 3D cinema there may be health related effects in the use of the same glasses by multiple viewers. The ITU-R is asking the World Health Organization for information on all these issues.*

What can the EBU do to help?

The EBU technology group will continue to monitor market developments and those in the production, distribution and consumer domains. In particular, the following areas will be covered:

- 3D content availability from the industry at large (Blu-ray, gaming, acquisitions);
- Production technology, training and operational guidelines;
- Distribution and consumer technologies;
- Recommendations to standards bodies;
- Representing PSBs' positions;
- Monitoring and disseminating information about the physiological effects of 3D viewing.

---

[1] See for example, http://www.telegraph.co.uk/science/steve-jones/7451130/When-watching-3D-can-fall-flat.html and http://www.settheory.com/stereo_blindness_test.html.

1. 3D history

What we currently think of as 3D is more accurately described as 'stereoscopy'. A true 3D image would allow you to see around objects in the picture. Also, as you moved, the image would look different viewed from different angles.

Stereo TV, on the other hand, follows 2D television rules i.e. the image follows you around the room and you can't look around objects. Stereo is more about providing additional depth rather than an additional dimension.

Stereoscopy experiments began in the 19th century, starting with still images but rapidly following the movies into the early cinema.

Stereo camera rigs were patented around 1900 and the earliest confirmed 3D film (*The Power of Love*) was shown in the Ambassador Hotel Theatre, Los Angeles, in September 1922!

Over the past 90 years, 3D has come and gone. After each decline there have been various attempts to revive the technology.

The 1950s were described as the golden age of 3D with the now-infamous '*House of Wax*' released in April 1953 with stereo sound!

The 1980s saw a run of 'Part III' films with the addition of a "-D" at the end of the title (e.g. '*Jaws Part IIID*'). Each revival was usually the result of a technical advance or a technique that seemed to make 3D better or more compelling, but it was never enough to catch a mass sustainable market.

2. Current situation

The latest revival of 3D is slightly different. It is being driven by a combination of factors:

- a run of very good 3D animations in the cinema that has whetted the public's appetite;
- the availability of affordable domestic flat screens capable of displaying several different 3D formats (prices starting from around 1400 €);
- in the near future, all HD Ready displays will be 3D capable;
- 3D content (feature films, gaming, US series) will be available direct to home via 3D Blu-ray disks and high speed internet (HBBTV, VoD, etc.);
- consumer cameras (stills and video cameras) are available that interface directly with the 3D displays.

3. Stereoscopy for Television

3.1 Acquisition

Traditional 2D camera positions for an event may not be good enough for 3D. If the cameras are too far away from the action, for example, the 3D depth effect is lost.

Mini Camera
Side by Side Rig

2 eyes = 2 cameras. But it's not quite as simple as strapping two cameras side-by-side.

The theory suggests the centres of the lenses should be about the same distance apart as human eyes. It's this distance that makes the depth 'seem' about right. Cameras are placed on a rig in different ways.

Only mini-cameras with small lenses are narrow enough for **side-by-side** rigs.

Studio cameras and field camcorders with high quality lenses are usually too large for side-by-side mounting (the lenses would be too far apart). Larger cameras therefore use 3D **mirror** rigs.

Full Size
Mirror Rig

The most common mirror rigs mount the cameras at 90° to each other with one looking through a 45° mirror.

Good mirror rigs can cost over 30 k€ and they are big!

Notes:

- *Lenses are matched as closely as possible for left eye and right eye cameras.*
- *Cameras are aligned and tracked as closely as possible.*
- *Zooms do not necessarily work well and some camera moves destroy the 3D illusion.*

Camera Positions for 2D and 3D coverage of football

Figure 1a: Camera positions for 2D production

Figure 1b: Additional cameras required for 3D production
(courtesy of HBS [http://www.hbs.tv/orientation])

The additional cost of producing 3D content is difficult to calculate accurately.

Until there is a lot more experience of 3D acquisition, using one (typically the left eye) view of a 3D production for 2D HDTV is normally not advisable. In consequence of this, events for 2D HD and 3D will require two production crews.

Studio infrastructures based on 3G-SDI are being developed and can be used both for 3D and for next generation HDTV (1080p/50).

3D single camera productions usually need a longer time to set up and light and they may in fact need separate 2D and 3D cameras on the same shots if they are to be used for 2D and 3D services.

Each programme type will need to be assessed for 3D suitability against cost and value to the audience. Estimates vary from 20 - 50% increase over 2D HD.

Is has been observed that there is now a trend for integrated 3D cameras (combining left and right image capture in one camera housing; various technologies are employed).

Several manufactures are developing production tools to make set-up and alignment for 3D productions easier.

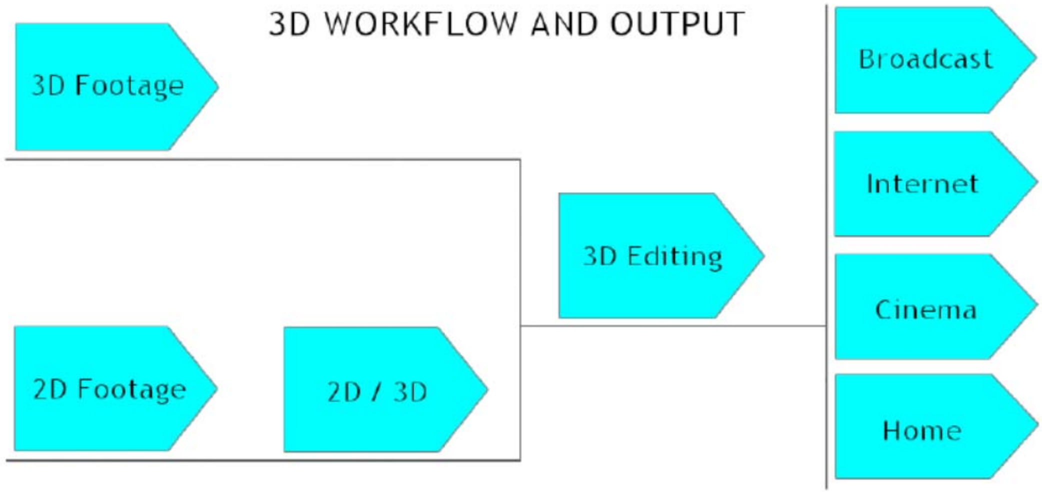


Figure 2: Simplified production route

3.2 Postproduction and Processing

If the cameras are set up properly, live and as-live programmes with simple cut editing are relatively straightforward. If the 3D depth of a programme needs to be adjusted in postproduction (with special 3D capable equipment and trained staff), this can take longer than editing the programme itself!

All major manufactures for postproduction equipment such as AVID, DVS, EVS, GV, Quantel and Sony are providing or updating their products to become 3D capable.

Getting things right in postproduction is a very time consuming and expensive thing to do. The shortage of skilled 3D editors will only add to the time and expense.

Storage requirements for 3D are double that of 2D and two images have to be moved around existing infrastructure and also archived.

All this will add to the cost and time taken to postproduce a 3D programme. Current thinking indicates that costs will increase by 25 - 30% over 2D programme costs.

3D postproduction costs will nevertheless fall as 3D options become standard in equipment.

Computer animation is ideal for 3D - hence the recent glut of 3D animated feature films. The rendering and postproduction computing power already required for (2D) animation is quite capable of handling the 3D element.

Animation has one other advantage as far as 3D is concerned; as everything in a computer animation is synthesised, if anything looks wrong, it can be changed to look right! The perspective or horizon or any part of the image can be adjusted until the 3D works properly.

3.3 2D to 3D conversion

It is technical possible to generate low quality 3D images from 2D images.

This process can be used where 3D cameras were not available for acquisition or for archived content to be included in 3D programmes.

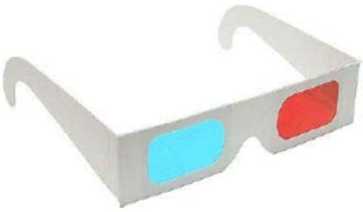
Extreme caution should be used for converting 2D to 3D.

3.4 Displays and glasses

How can 3D work when it's displayed on a 2D screen? There are currently four primary technologies:

3.4.1 Colour separation

Anaglyph and Colorcode (and other colour variants) work by showing both left and right images at the same time on screen, the left eye image in one colour and the right eye image in another. This was the 'original' 3D technology used in cinemas in the past ('*House of Wax*' etc.).



Anaglyph Glasses

Colorcode Glasses

The viewer uses coloured glasses to make sure that the images go to the appropriate eye.

**Figure 4:
Red/Cyan anaglyph image as it is transmitted and appears on all TVs**

Whilst this is an inexpensive technology that works on any available display, the resulting 3D image quality is very unsatisfactory. It is no longer competitive in face of newer technologies and it is possibly the reason why 3D was never accepted in the past.

Consequently, EBU Members are advised not to use this technology.

3.4.2 Polarised / Passive

This technology demands a new 3D TV display with a polarised screen. The left and right eye images are lined up behind the polarising screen and each image is given a different polarisation.

To see the 3D image, polarised glasses are needed. These are similar to polarised sun glasses but the lens of each eye has a different polarisation to make sure the correct image goes to each eye.

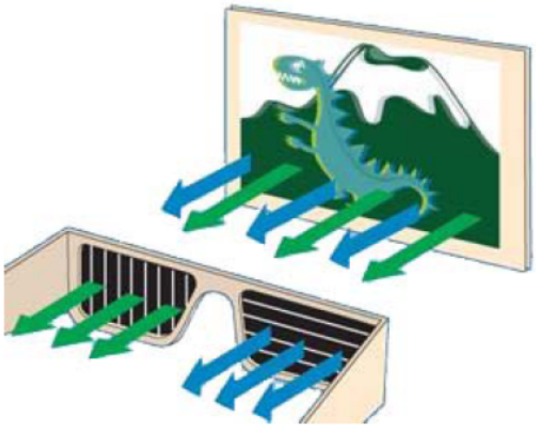
The glasses include no active electronic elements and are therefore termed 'passive'.

Advantages

- Good 3D;
- No colour or picture distortion as with colour separation;
- Glasses are very cheap;
- Used in displays for the professional market.

Disadvantages

- Requires a new 3D TV display;
- Both images are reduced resolution (no longer HD resolution);
- Technology in the display is more complex than for displays with shutter glasses (see below), thus they are not very common in the consumer display market;
- Variants of polarisation technology (e.g. linear orthogonal, circular) mean that glasses may not be compatible between different displays.



3.4.3 Shuttered / Active

Shuttered technology requires two active components - the screen and the glasses.

Shuttered LCD glasses are controlled by an infra-red signal sent from the TV. The left and right images are displayed alternately on the screen at a high frame rate (100 or more frames per second). When the left eye image is on screen the right eye lens of the glasses is made opaque and vice-versa.

Advantages

- Good 3D;
- Full resolution images;
- This is the primary format chosen by display manufactures for the new consumer 3D TVs.

Disadvantages

- Glasses are expensive e.g. from 80 € a pair;
- The Infra Red signal has not been standardised (e.g. glasses of one display will not work on another manufacture's display);
- Interference of the infra red signal can cause flickering.

3.4.4 3D TV with no glasses

The screen is covered with tiny lenses, arranged to send viewing zones of left/right images to the viewer.

The lenses direct the left/right images out of the screen in zones. If you sit in a zone at the correct distance and angle you see 3D - if you move out of a zone you lose the 3D image.

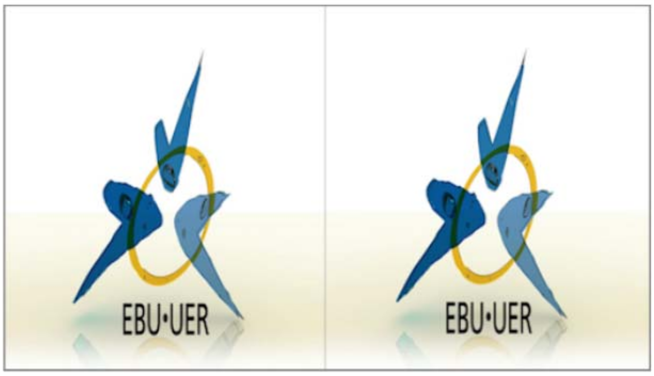
Lenticular screen technology is still very new. Displays for the consumer market are expected from about 2014 onwards.

4. Distribution and receiver

4.1 Current options

This involves using current HDTV distribution technologies and current set top boxes (STBs) in the home, but a new display is required.

The left and right signals are combined in a HDTV frame with the consequence that both images are reduced in resolution. This is called ***Frame Compatible*** 3D. Various ways of combining the signals are possible, the most prominent are:



Side-by-Side

Top and Bottom

The STBs pass these images through and it is the new 3D displays that stretch the reduced resolution left and right eye images to full frame to create the 3D image.

Note: *Current 2D HDTV displays (not new, 3D displays) would show a side-by-side (or top-bottom) image as illustrated in the above figure, which is of course unusable for the viewer.*

4.2 Future technologies in distribution and production

The limitations in 3D image quality inherent with Frame Compatible technology have triggered several working initiatives in the standards organisations such as DVB (for distribution) and the SMPTE (for production). Some of the aims of these activities are:

- Improved 3D image quality;
- More features such as depth control on the receiver or display (consumer side);

- 2D HD backwards compatibility to avoid the need for simulcasting, called ***Service Compatible*** 3D;
- 3D subtitling;
- Contribution links;
- Broadcast equipment interoperability standards.

5. Glossary of important terms

These are some of the more common terms used in 3D production; many more are used by experts.

Convergence:	Is where an object within a frame appears as a single image (both left and right images sit perfectly on top of each other). It is the position or plane of the TV screen and objects in front will appear to be out of the screen and objects behind will appear to be into the screen.
Depth budget:	This is the amount of 3D that is established for the 3D production. Typically it is referred to as a percentage. For some 3D productions of live events, depth budgets of less than 4% are used, e.g. 2.5% into the screen 0.5% out of the screen have been used.
Disparity:	Distance of the same object between the left and right eye view.
Inter Ocular / Inter Axial:	Commonly termed IO this refers to the physical distance between the centre of 2 frames, or loosely how far the cameras are apart. Note: it is the relationship between the Convergence and Inter Axial that determines the amount of 3D.
Positive & Negative Parallax: ("Inies" and "Outies")	Positive parallax means the object appears behind the screen, hence the term "Inies". Negative parallax means the object appears in front of the screen, hence the term "outies". Zero parallax means the object appears on the screen and the left and right eye images are exactly overlaid (exactly like a 2D image)
S3D:	Stereoscopic 3D
Stereographer:	New role in a 3D production team. He or She takes responsibility for the quality of the 3D.
Z-axis:	The axis from front to back of a 3D image, i.e. into and out of the screen.

Annex 9

A preliminary set of possible performance requirements for a 3DTV broadcasting service

This Annex is a collation of candidate performance requirements for a 3DTV broadcasting service that have been identified in contributions submitted to WP 6C since October 2008. As far as possible, each is presented in the form in which it appeared in the original contribution, and the source contribution is indicated. They are listed in order of the contributions in which they are mentioned.

It should not be assumed that this list is comprehensive, or that each would necessarily be applicable to all possible types of 3DTV services.

a) The goal of the 3DTV viewing experience should be to create the illusion of a real environment, which can be watched for an indefinite period of time by an audience with normal visual acuity (Document 6C/128).

b) The quality of the 3DTV service should be established by two principal parameters: sensation of reality and comfortableness or ease of viewing (Document 6C/128).

c) In the case of a 3DTV emission system that is intended to allow for the compatible reception of conventional SDTV, EDTV or HDTV digital signals, no perceptible degradation of the conventional television presentation should be caused by the additional signal coding required to convey the extra information needed for 3DTV (Document 6C/128).

d) The viewing experience provided by a new 3DTV service should represent a major improvement as compared to the experience provided by HDTV broadcasting at its best quality level (Document 6C/155).

e) A new 3DTV service should be implemented in a way to match the viewing habits that will likely be typical of television audiences at the time when the service will be introduced (Document 6C/155).

f) A single set of image and sound specifications should be recommended for the new 3DTV service, unless it is essential to choose more than one set of specifications in order to reflect unavoidable regional constraints (Document 6C/298).

g) It would be desirable, albeit not imperative, to preserve a certain degree of backward compatibility of the new 3DTV service with existing consumer receivers (Document 6C/298).

h) The image and sound system parameters of each new television broadcasting application should be optimized on the basis of the performance requirements specific to the 3DTV service, even if this requires departing from technical solutions adopted for other applications (Document 6C/298).

j) Specifications should be recommended for a new 3DTV broadcasting service after it has been demonstrated through life-size simulations and tests that it offers new attractive features, and that broadcasting audiences that have been exposed to it, have appreciated its features and have expressed their interest to gain access to it, purchasing the necessary consumer equipment (Document 6C/298).

k) The issue of visual fatigue and other possible health hazards related to extended viewing of 3DTV programmes based on presentation methods currently available or proposed, should be resolved prior to developing Recommendations on technical specifications for a future 3DTV broadcasting service (Document 6C/299).