International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Report ITU-R BT.2160

(11/2009)

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, 2010

REPORT ITU-R BT.2160

Features of three-dimensional television video systems for broadcasting

(2009)

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 has announced its intention to introduce a stereoscopic 3DTV channel in the United Kingdom during 2010. Several consumer electronics manufacturers have also announced their intention to introduce 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 is creating expectations of the possibly 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

---

[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] Currently available Blu-ray and DVD packaged media use a rudimentary form of stereoscopic television, referred to later as a Level 1 system. The expectations are that a more sophisticated system will soon be available.

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 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.

2 Background to possible 3DTV systems

The fundamental means by which a 3DTV broadcast system today is capable of enhancing users' visual experience of three-dimensionality, compared to the broadcast of high-definition television (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 provides 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 Fig. 1 below. 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

(1) MPEG-4 AVC Stereo High Profile – a subset (to be published) of MPEG Multiview Video Coding (MVC), annex H to ITU-T Recommendation H.264.

(2) It is understood that ISO/IEC JTC/SC29/WG11 intends to address this form of extensoion to MVC.

(3) Annex G to ITU-T Recommendation H.264.

Report 2160-01

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). It is understood that BSkyB in the UK will probably use the SbS method (2x1080i/960) subject to discussions on licence issues. 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.

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 allows 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. 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 ac 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 autosteroscopic 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.

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 4. 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.

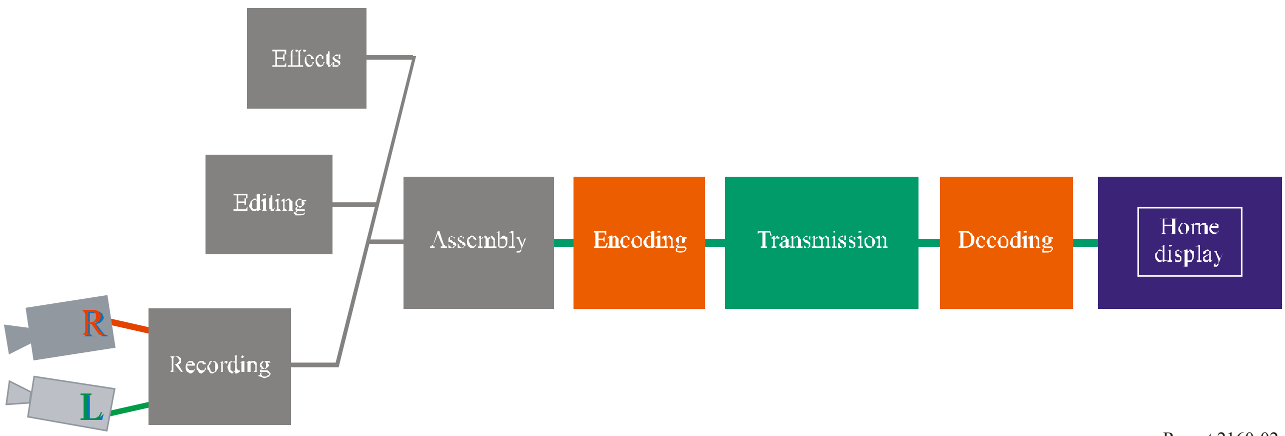
Last, but not least, independent testing using a standardized testing methodology is needed in order to accurately quantify how much extra bit rate would be needed, using a range of representative 3DTV source material.

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 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.

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. Live 3DTV production presents particular challenges.

Factors that affect 3D viewing comfort 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 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.

9 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.

10 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.

11 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 respondees considered it very important/essential for a 3DTV system to have the same format as packaged media (e.g. HDTV capacity discs).

12 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.

13 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;
– 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.

It is also understood that amendment to ISO/IEC 14496-10 are to include a spatially interleaved frame supplemental enhancement information (SEI).message to signal the type of interleaving in a frame-based scheme.

2 3DTV – Network of Excellence

See www.3dtv-research.org.

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

3D4You is a project funded by the European Union under the ICT (information and communication technologies) 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/index.php.

4 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.

5 SMPTE

In August 2008, SMPTE formed a Task Force on 3D to the Home. This completed its work with publication of final report in March 2009. The report is available to SMPTE members, and also from http://store.smpte.org/product-p/tf3d.htm. The Recommendations include standardizing a "3D Home Master", based on the requirements developed by the Task Force. SMPTE is starting work on the specifications based on 1 920 × 1 080 pixel resolution at 60 frames per second and per eye.

6 The Digital Video Broadcasting Project (DVB)

Technical work in DVB is driven by commercial requirements. Following completion of a study mission to investigate the possible need for 3D activities, it has been decided to establish an ad-hoc group to consider the commercial requirements for stereoscopic 3DTV.

7 The Blu-ray Disc Association (BDA)

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

8 HDMI Licensing, LLC, has announced the release of HDMI specification 1.4.

See: http://www.hdmi.org/press/press_release.aspx?prid=101.

9 Consumer Electronics Association (CEA)

The 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.

10 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.

11 Association of Radio Industries and Businesses (ARIB)

The ARIB has established a working group for researching 3DTV broadcasting in 2008.

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

12 Ultra-realistic communications forum (URCF)

The 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/.

13 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/.

14 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/.

15 Japanese Ergonomics National Committee (JENC)

JENC is in charge of the national preparation for ISO.TC159.

See: http://ergonomics.jp/jenc/index.html.

Annex 2

Historical background on the development of stereoscopic and 3D television systems. To include Document 6C/92 – The present state of three-dimensional (3D) TV broadcasting studies in the Russian Federation see:

6C_92.doc

Annex 3

Introduction to free viewpoint television: Annex 1 to Annex 6 to Document 6C/69 see:

069N01e.doc 069N02e.doc 069N03e.doc 069N04e.doc 069N05e.doc 069N06e.doc