BBC

Research White Paper

WHP 173

January 2009

The Challenges of Three-Dimensional Television

S Jolly, M Armstrong *and* R Salmon

BRITISH BROADCASTING CORPORATION

BBC Research & Development
White Paper WHP 173

The Challenges of Three-Dimensional Television

Stephen Jolly, Mike Armstrong and Richard Salmon

Abstract

The recent resurgence of interest in the stereoscopic cinema and the increasing availability to the consumer of stereoscopic televisions and computer displays are leading broadcasters to consider, once again, the feasibility of stereoscopic broadcasting. High Definition Television is now widely deployed, and the R&D departments of broadcasters, consumer electronics manufacturers and other organisations are starting to plan future enhancements to the experience of television. Improving the perception of depth via stereoscopy is a strong candidate technology.

In this paper we will consider the challenges associated with the production, transmission and display of different forms of "three-dimensional" television. We will explore options available to a broadcaster wishing to start a 3D service using the technologies available at the present time, and consider how they could be improved to enable many more television programmes to be recorded and transmitted in a 3D-compatible form, paying particular attention to scenarios such as live broadcasting, where the workflows developed for the stereoscopic cinema are inapplicable.

We will also consider the opportunities available for broadcasters to reach audiences with "three-dimensional" content via other media in the near future: for example, distributing content via the existing stereoscopic cinema network, or over the Internet to owners of stereoscopic computer displays.

A substantially similar paper to this one was presented at Stereoscopic Displays and Applications XX, and appears in the proceedings of that conference [1].

Additional key words: 3DTV, Beyond, HD, SDA2009

White Papers are distributed freely on request.

Authorisation of the Head of Research is required for publication.

BBC Research & Development
White Paper WHP 173

The Challenges of Three-Dimensional Television

Stephen Jolly, Mike Armstrong and Richard Salmon

1 Introduction

Stereoscopic television was first demonstrated by Baird in August 1928 [2], some years before the BBC started its first regular television broadcasts [3]. His system incorporated a prismatic stereoscope for a single viewer, which would have been a barrier to its adoption unless, as Tiltman hints [2], "free viewing" of stereoscopic television pictures presented side-by-side were to have become widely accepted.

The history of stereoscopic television between that first demonstration and the present day has consisted mostly of novelty programmes and a few experiments. The introduction of colour television in the 1950s and 60s permitted the broadcast of anaglyphic programmes, and indeed more than 15 million pairs of anaglyph glasses were sold in Germany and Austria to watch several transmissions in early 1982, but public interest proved to be transitory [4]. There have also been a small number of "3D" broadcasts that make use of the Pulfrich effect. The BBC transmitted a special episode of "Doctor Who" in this format during the 1993 broadcast of its annual "Children in Need" telethon. Viewing glasses were sold in newsagents, with the profits going to charity.

Where the development of the creative stereoscopic medium is concerned, the baton has been taken up in the world of cinema. A thorough survey of that work in the USA and Russia up to the 1980s can be found in Lipton's well-known book [5]. In the UK, a significant amount of work was done on stereoscopic production by Raymond and Nigel Spottiswoode and Charles W. Smith in the early 1950s. The work was undertaken for the British Film Institute and the Festival of Britain [6, 7]. In the 14 months running up to the Festival of Britain they produced the first films with both stereoscopic images and "4-channel stereophonic" sound. Charles W. Smith continues working on stereoscopy and worked with Andrew Drumbreck, then of Harwell Laboratories on the design of stereoscopic equipment for remote handling in the nuclear industry [8]. This collaboration then formed the basis for further work on broadcast television [9], within the European ACTS MIRAGE project in the late 1990s. In June 1996 what was claimed to be the first ever broadcast quality stereoscopic television programme, "Eye to Eye", was made at Anglia Television as part of that project [10]. More recently, the European ATTEST project [11] demonstrated the benefits of "2D+depth" representations of 3D video content, in which the third dimension is directly represented by a depth map. A successor project, "3D4YOU" (in which the BBC is participating) is currently researching the best way to capture material in this and similar formats and to demonstrate a full 3DTV broadcast chain [12].

The recent interest in and financial success reported by digital stereoscopic cinema [13] has led many within the broadcast industry to review the possibilities of stereoscopic (and other forms of 3D) television. A number of broadcasters and other organisations including the BBC [14] have experimented with closed tests of transmitting live events in stereo – usually sports matches, but with the exception of Japanese cable and satellite broadcaster BS-11 [15], who have been offering several short broadcasts of stereoscopic material daily since December 2007, little if any content has made it to the public airwaves as yet. However, broadcasters should obtain experience with 3D production well before considering the launch of a 3D television service: this is what happened with High Definition Television (HDTV), where the BBC wanted to co-produce high-budget series with partners in markets where HD services had already launched (and indeed to sell programmes to those markets, and to repurpose programmes for cinematic release).

In this paper we assume that stereoscopic television will develop as an enhancement to HDTV, as part of a continuing long-term evolution of television standards towards a means of recording, transmitting and displaying images that are indistinguishable from reality. Multiview autostereoscopic display technologies already astonish their viewers, but at present these lack a general-purpose solution for capturing material in an appropriate format, and hence this paper will concentrate on stereoscopy.

This paper assumes an understanding of the fundamental principles of stereoscopy. Some good introductions to this field can be found on the Internet – see [16, 17, 18] for example.

2 A Comparison of Production for the Cinema and for Television

The equipment and working practices of digital film-making appear very familiar to the broadcaster, who uses them more or less modified to produce "drama" content for television. Although stereoscopic content of this kind may be broadcast on television at some point in the future (whether made for television or originally shown in cinemas), the general requirements of a stereoscopic television service are far more complicated, however.

A film or drama production is characterised by considerable control over the environment in which it is made. Actors can be asked to repeat lines or scenes, studios can be filled with an environment that matches exactly the creative vision of the production team, and "effects" such as the use of CGI elements or miniatures can be used to present the viewer with images that never existed in the real world.

In most television productions, that control is usually reduced, often severely. Broadcasts of live events such as sports are extremely restricted in terms of camera locations, and the action is rarely under directorial control. Live events preclude the use of post-production to repair or augment the material captured on camera. Even for pre-recorded programmes, post-production beyond basic editing, adding titles and graphics, a sound track and perhaps some simple video effects and compositing is rare. There is a much greater need to get things right "in the camera" which has considerable implications for today's stereoscopic camera rigs and workflows, which often rely heavily on post-production.

In general, the most significant difference between producing mainstream content for the cinema and for television is the budgets involved. It is reasonable to suppose that any commercial stereoscopic television service will be limited initially to those "premium" genres where viewers are prepared to pay to cover the additional costs of stereo production and transmission, such as sport and films, but even for such a limited service, there will be pressure to keep costs down. In a multi-genre public service broadcaster such as the BBC, it would be reasonable to expect stereoscopic production to focus initially on content where the medium adds the most impact and where more of the money can be recouped in cinematic presentations and resale for broadcast in other markets. Another driver for stereoscopic production may be co-productions with partners in markets where uptake of stereoscopic television takes place sooner, as happened with HDTV. Again this implies a focus on premium genres; perhaps drama and natural history. The intention for a public service broadcaster must be ultimately to cover all genres where stereoscopy would deliver clear benefits to the viewer in terms of both entertainment and the explanation of factual information, however. This is where the real challenge lies: these other genres are characterised by lower budgets and shorter lead times, which implies substantial advances in the technology of stereoscopic production: it must become cheaper, simpler, more reliable and more automated if it is ever to make serious inroads into mainstream television.

3 Artistic Considerations

The stereoscopic cinema is evolving a grammar distinct from that of 2D production, and there is every reason to believe that a similar process will take place if stereoscopic television starts being seriously developed as an independent medium. Many staples of 2D production simply do not work stereoscopically: for example, false perspective, "over the shoulder" shots during dialogue and the use of matte paintings. Others will require experimentation: there are dissenting views at present on the use of dissolves to transition between shots in stereoscopic video for example; indeed many transitions and other video effects are not obviously appropriate for stereoscopic material. This issue is discussed in more detail in section 5.

The highly constrained camera positions that are a feature of many television productions introduce problems. If telephoto lens magnifications are required to frame a distant subject, the stereographer must find a compromise where the camera interaxial is concerned, as widening it to increase the range of disparities associated with that subject and improve its three-dimensionality has well-known drawbacks. Doing so will result in "dwarfism", and also has the potential for causing extreme viewer discomfort if (for example) a foreground object passes in front of the camera. The optimal compromise (in this and other matters) will probably evolve as audiences for stereoscopic content become more aware of the limitations of the medium, in the same way as film audiences around the start of the twentieth century learned to accept the "cut" as a way to build rich narratives within that medium's limitations.

There are techniques that are commonplace in monoscopic television and cinema that are still considered hard to replicate in three dimensions. Zooms during shots are the canonical example, since the "telecentricity" property of lenses—the extent to which the image moves perpendicular to the lens optical axis during a change of focal length—is not tightly controlled by lens designers at present, and two randomly selected lenses are unlikely to be well-matched. Mechanical shifting of one camera or real-time electronic shifting of its image as the lenses are zoomed is a more likely option at present – the best-known implementation of this approach is by 3ality [19]. There is a significant calibration overhead associated with this, however, and the technique is liable to run into limitations for scenes containing fast motion. Manual correction in post-production is possible for non-live programmes, which removes the requirement for calibration but is subject to the same motion limitations. If stereoscopic production becomes widespread, tighter control of telecentricity in the lens design process or the availability for sale or purchase of matched lens pairs (or even combined lenses, similar perhaps to the prototype developed by Canon for their XL1 camcorder [20]) could eliminate the issue. A. Drumbreck reports [21] that for experiments performed by the UK's Independent Television Commission in the 1990s, two lenses were selected for their similarity from a large sample. This is straying a little far from artistic considerations, however. The recent 3ality-produced feature "U2-3D" made sparing use of zooms, but aside from that it remains on the list of shooting techniques whose applicability to stereoscopy remains to be fully explored.

When we talk about stereoscopic television, we must bear in mind that for some time to come (and perhaps indefinitely), not all television viewers will have the equipment necessary to receive stereo broadcasts. For some kinds of service this is not a problem – live sports broadcasting via satellite could well be economically viable in some markets as a service independent from 2D production, for example, with its own production facilities and transmission bandwidth. Most television services are unable to command high subscription revenues, however. It is therefore fair to say that if stereoscopic television starts to move into the mainstream, a key part of making it affordable will be to take advantage of the backwards compatibility that stereoscopy offers with monoscopy and plan to use one eye view of the stereoscopic programme a 2D programme in its own right. This has many technological implications, and one important artistic one: programmes will have to look good in both two and three dimensions – without any re-editing, where live programmes are concerned.

The production grammar changes that will be needed imply a very significant requirement for retraining: the changes are likely to be far more challenging than those associated with the adoption of High Definition television, itself considered a major change within the television industry. A strong case can be made for extensive testing and experimentation at an early stage, to identify as many of the new restrictions as possible that will be imposed on stereoscopic production by the nature of the medium and the limitations of early equipment. Avoiding these and finding out how to best exploit the medium to deliver the best possible experience to the viewer will be a significant challenge that must be addressed alongside the engineering research required to determine the technical aspects of a stereoscopic broadcasting system. Broadcasters will have to invest time and money in understanding what makes good stereoscopic material before committing larger sums of money to stereoscopic productions.

It is early to be considering what production techniques and workflows might evolve as stereoscopic television develops, but techniques such the "slaving" of convergence to the lenses' focal distance as promoted by Cameron [22] may help reduce the number of parameters that the camera operator must take active responsibility for. The issues associated with the use of telephoto lenses and zooming may lead to more use of steadicam – KUK-film in Germany [23] have achieved good results using this technique as part of a mixed approach to stereoscopic production, although it would be liable to push budgets up a little.

At present, the workflows demonstrated for stereoscopic production using broadcast equipment have been labour-intensive compared to monoscopic productions of similar scale. For the vast majority of television productions, for example, the budget will not stretch to a separate stereographer – in general that role will be filled by a multi-tasking member of the production team. Stereoscopic workflows will have to evolve alongside improvements to capture and post-processing technology to find the right compromise between viewer comfort, quality of experience and cost-effectiveness.

4 Stereoscopic Capture and Storage

As stereoscopic film-makers have already discovered, the digital video cameras developed for broadcasting and digital film-making can be readily adapted to stereoscopic capture, given an external source of genlock (and perhaps timecode). A desire to keep the camera interaxial similar to that of the range of human interoculars while retaining the high image quality associated with professional cameras for HD and digital film-making (for which high-quality lenses almost invariably have diameters greater than this range) has led to the widespread use of mirror rigs for camera mounting (see Fig. 1). For post-produced material, correction of the artefacts thus introduced is perfectly feasible, and real-time correction using custom hardware or programmable Digital Video Effects (DVE) units has been demonstrated. Mirror rigs are now available for hire (by themselves or with an operator) and purchase, which is a key step in the uptake of any film or television capture technology.

A variety of miniature HD cameras have been developed in recent years, and in theory there is no reason why such devices could not be used for stereoscopic capture. The advantage of these miniature cameras is of course that they are often physically small enough to achieve an interaxial similar to the human interocular without requiring the use of mirrors. The small sensors facilitate the design of small lenses, which are another clear prerequisite for this kind of capture. From a television perspective they generally possess two characteristics that limit their usefulness, however. Firstly, they tend to use significantly smaller sensors than the "2/3 inch" chips found in full-size broadcast cameras. This implies shorter focal-length lenses for the same angle of view, and hence smaller aperture diameters for the same f number, which leads to visible diffraction effects at small apertures and HD resolutions. The small sensors also imply smaller pixels and hence lower sensitivity, which leads to noisier images if lighting levels are not altered. Image noise masks detail and leads to spurious high spatial frequencies that visibly reduce the performance of the video codecs used for transmission – indeed the BBC advises HD productions against the use of cameras with sensors smaller than the "1/2 inch" standard [24] for this reason. The second characteristic

Figure 1: A photograph of the P+S Technik mirror rig, shown here equipped with Silicon Imaging SI-2K Mini camera heads (right and, partially concealed, on top). Also visible are a P+S Technik 3d sync box (mounted top right) powered by a camera battery (resting on top of the rig, in front of the top camera), which provides timing information to synchronise the cameras.

that can limit these cameras' usefulness in television productions is their tendency to use Bayer-pattern sensors with native resolutions no higher than 1920x1080. The luminance resolution of such a sensor is considerably less than this: around 720 lines if a good interpolation algorithm is used. (720-line cameras are generally considered to be Standard Definition by the BBC [24], although no explicit guidance on the use of cameras with Bayer-pattern sensors is currently published). Obviously this becomes even more of an issue in a stereoscopic production where further resolution may be lost due to keystone correction and/or image shifts. In addition, it has been reported [25] that it is extremely hard to process the information from cameras with Bayer-pattern sensors into a viewable stereoscopic image (a process known as "debayering") without introducing visible artefacts.

Overall, most of the equipment required for conventional stereoscopic production is now available for purchase or hire. Cameras and rigs have been mentioned above. Timecode and genlock sources are readily available. For recording purposes, there are a wide range of digital recording systems available for purchase and hire, both tape- and file-based, which increasingly have dedicated support for stereoscopic material. The question of monitoring is harder to address, though. Viewing the camera output, even stereoscopically, is not a hard problem, but the issue of monitoring for a stereoscopic production goes beyond this. Firstly, the majority of scenes will require at least a horizontal image shift or a keystone correction operation to be performed before the material can be viewed comfortably. By mid-2009 it is probable that one or more simple video processing units that can perform this kind of manipulation will be on the market. However, due to the probable mismatch between the size of the on-set monitor and the display size (or range of display sizes) for which the material is intended, it is unlikely that an image optimised for the former will look good (or even be acceptable) on the latter, even if the stereographer is sufficiently experienced to judge how the image will appear to the eyes of an untrained viewer. To take the guesswork out, stereographers have historically set the cameras up based on careful measurement of the scene (which is time-consuming) or by calculating the permissible disparity, scaling it to account for the monitor size mismatch described above, and comparing it to the amount of horizontal image movement required to move the screen plane from the near point to the far point of the stereoscopic image. This can be done either by mechanically shifting one of the screens in a beam-splitter based stereoscopic monitor, or electronically shifting one of the images with a DVE unit.

Some organisations have gone a step further. Software-based disparity estimation has been an area of academic research for decades; it finds applications in machine vision and cartography, amongst other areas. Recently, the technique has started to be applied to various aspects of stereoscopic video and film production. It has an obvious application to stereoscopic monitoring: if it can be performed accurately, the process of ensuring that the range of disparities will not exceed a comfortable viewing range can be automated. More than one company has developed such a production tool, generally for internal use, displaying a 2D image augmented with this estimated disparity information. Ultimately, such tools could greatly simplify the workflow of stereoscopic production, perhaps allowing camera operators or stereographers to select two key distances in the scene (eg far plane, screen plane) on a touch-screen monitor and have the camera setup adjusted automatically.

Over time, camera systems will evolve to reduce assembly and set-up time, and improve operability. In the short term, camera bodies with the option of a plug-in module containing a second optical block and sensor assembly would be a sensible way to build a single-body 2D/3D camera compatible with existing stereoscopic rigs. Built-in automatic or semi-automated calibration for exposure and colour balancing (including the removal of the graduated colour cast introduced by mirror rigs) would reduce the amount of expensive post-production time required, and the external genlock devices currently used to synchronise pairs of cameras would be redundant. In the longer term, disparity estimation and view interpolation techniques may mature to the point where they can simulate acceptably high-quality views from virtual camera positions between the two cameras of a stereo pair – perhaps even in real time. This would enable, amongst other things, the development of side-by-side camera rigs without lens-size restrictions and with no moving parts other than the lens mechanisms, simplifying the workflow of shooting stereoscopically still further in shooting environments where the view interpolation produced good results.

Although capturing stereoscopic content is now practical, there is some distance to go before it can be considered convenient. Devices capable of synchronously recording two streams of HD video are a good start (two streams of 1080p50/60 HD would be even better from a television perspective), but a system that relies on human operators to consistently connect the cables up the same way around, and then keep the left and right-eye video sequences associated correctly throughout the post-production and transmission process is arguably insufficiently mature. The recording systems of the future will handle this association of stereoscopic streams automatically. The important thing is to standardise interfaces: a standard way of conveying twin streams of 1080i25 video down a single 3Gbps HD-SDI link, for example, and stereoscopic profiles for AAF, the Advanced Authoring Format that has been standardised as a format for storing video files and their associated metadata [26], and MXF, the related Media Exchange Format for transferring file-based media between production and post-production tools [27].

Capture of stereoscopic metadata during the shoot is extremely important. Information regarding the camera interaxial and convergence (if any), the camera position in the rig (which eye view it corresponds to, and what reflections it may have been subjected to by a beam-splitter), the display size targeted by the stereographer, the horizontal parallax at which the subject of the scene is located, *etc* should be recorded alongside the video. A standardised format for storing this information in file formats like AAF and MXF would permit editing software to make a first attempt at rectifying the image, potentially reducing the time spent in post-production.

5 Stereoscopic Post-Production

Offline editing of stereoscopic material is already officially supported in a small number of editing systems and software packages, and is available as a service from post-production facilities. It is safe to assume that basic support for stereoscopic content (the ability to keep the video corresponding to the two stereo views associated in the timeline, and treated as a single entity for most editing operations) will be added to non-linear editing software at all price levels as and when existing customers for that software start to experiment with stereoscopy in sufficient numbers. What will be interesting is the evolution of software tools within those packages to assist with the stereoscopic editing process (eg automatic image realignment (or "rectification") and colour matching of the two camera views of a shot), and the extent to which those facilities can be implemented for the more challenging environment of live production.

Television productions that make extensive use of archive footage are commonplace, and the incorporation of footage into stereoscopic productions from archives that are overwhelmingly monoscopic will be an interesting challenge. Conversion of the material to stereo by adding a second-eye view is likely to be prohibitively expensive for all but the highest-budget programmes. It should be noted that unlike stereoscopic material, 2D footage can be placed at a different position and orientation within a monoscopic or stereoscopic scene by the use of a perspective projection. In practice, the treatment of 2D material is likely to depend on the context in which the material is shown, as material intended to be appreciated as historic will not suffer from being presented without depth. Other genres are likely to choose to make programmes containing a mixture of monoscopic and stereoscopic material deliberately, even if archive material is not to be used. News for example, would be likely to retain two-dimensional video for live interviews for as long as the editorial device whereby a presenter converses with a giant video wall remains popular. It may also be found that 2D video is perceived as "more authentic" by viewers – certainly a similar argument has been cited as a reason for news channels to upgrade studios to HD: to increase the difference in quality between studio-originated and location-originated material.

It is common for a broadcaster to incorporate many elements into the broadcast image other than the unprocessed footage from the television camera. On-screen "graphics" elements such as channel identifiers and scrolling text messages, transition effects such as dissolves, wipes, "turning page" effects and so on, and manipulation of the video image itself: scaling it, "rotating" it about various axes, etc. These will all require careful consideration to determine if and how they can be used in stereoscopic productions.

Where graphics are concerned, the most important rule is that they should never be rendered at a depth that is behind the part of the scene that they are overlaying. To a broadcaster who is used to overlaying a channel identifier in the same location on the screen for the entire duration of a programme, this has radical implications. Unless the use of negative parallax effects is outlawed (which seems unwise) or restricted to "safe" areas of the stereoscopic image (which seems impractical), automatic overlaying of graphics at the screen depth is not an option. Other possibilities exist, and the safest and easiest would seem to be to reserve an area at the top or bottom of the screen for graphics, cropping or scaling the video image appropriately. This solution would also work for closed captions, which are a legal requirement in some jurisdictions (including the UK) for a substantial fraction of programming, and differ from graphics in that they are applied by the television receiver. We leave the related issue of how (and if) to implement set-top box and television "on-screen displays" such as channel and volume indicators as an exercise for their manufacturers.

Scaling and other manipulation of video content is a different issue. Straightforward scaling (for example, placing a quarter-size video image in the corner of a screen containing text and other graphics content) is possible: the parallax range will be reduced accordingly, and it may be desirable in some circumstances to adjust the absolute parallax values with a horizontal image translation, to balance the depth of the video elements with that of the graphics. The reduction in parallax range associated with the downscaling ensures that this is possible without exceeding the parallax range of the original video. "Zooming in" digitally on a video image increases the parallax range, so is only possible if the original image had a relatively small range to start with, or if the zoomed area of the image excludes the areas of the original that contained the extreme disparities. A large number of more complicated video effects can be regarded as treating a two-dimensional video sequence as a texture to be mapped onto a shape (rectangular or otherwise) in 3D space, and these in general will not work with a stereoscopic image, as they break the geometric relationship between cameras and eyes that underlies stereoscopy. The well-known "page turn" transition effect is a good example of this, although there are many others.

Other forms of transition between sequences of video are similar again. Even simple cuts have issues associated with them, with stereoscopic theory and practical stereographers both promoting the importance of minimising the reconvergence of the viewer's eyes between cuts. This will be a particularly interesting challenge for television, where the subject of the viewer's attention is not always controlled as carefully as it is in cinematic productions. The effectiveness of "wipe" transitions between two stereoscopic scenes is similarly likely to depend on the content of the two scenes. If one or both scenes contain elements with negative parallaxes, the wipe might itself have to have a negative parallax associated with it, to avoid the same impression of objects appearing or vanishing first in the "wrong" eye that already leads to the use of a "forward window" in stereoscopy. The same applies to other transitions where a hard-edged boundary unmasks the new image. Dissolves are more problematic. A number of stereoscopic productions have experimented with them, but audience appreciation has not always been universal. It seems safe to say that more experimentation is required to determine the circumstances in which they work best. This also applies to the other transitions and video effects that involve the mixing of two video images.

In general, we should expect many of the conventions and assumptions behind two-dimensional video manipulation to be incompatible with stereoscopy. One obvious option which would enable attention-catching monoscopic effects and transitions to be retained would be to shift the video to a monoscopic representation before applying the effect. This would facilitate backwards compatibility with monoscopic viewing during the migration to stereoscopic television, since the 2D experience of a conservative stereoscopic production is likely to be unimpressive compared to competing 2D programming. Such a shift from stereoscopic to monoscopic representation might be disconcerting to the stereoscopic viewer, however, particularly if high-quality view interpolation (which would permit a fast yet smooth reduction of the displayed parallax to zero) is not available.

The fascinating thing to see will be what effects and techniques evolve that take advantage of the stereoscopic medium to offer new experiences that are unimpressive or simply unavailable in two dimensions.

6 Stereoscopic Display Considerations

We see stereoscopic television in the longer term as offering an experience to the viewer that is superior to HD. In the short-term, technologies may be adopted that deliver or display the stereoscopic image at the expense of some other aspect of the video quality, such as resolution or frame rate. Care must be taken when adopting a delivery format to avoid compromising the quality of the two-dimensional image that will be presented to a viewer without stereoscopic display equipment if that manner of backwards-compatibility is adopted. In any eventuality, the use of video compression for the delivery of stereoscopic television will still limit the quality of the image delivered to the home.

Television displays vary widely in size and display characteristics. The former has a well-known impact on the display of stereoscopic material, and presents a particular challenge to the broadcaster. For subscribers to a "premium" service, particularly for the first mover in a given market, the availability of compatible displays could be controlled or restricted by the broadcaster. A mature market for "3D" televisions is likely to contain a far greater range of display sizes and types, with the influence of broadcasters likely to be limited to collaboration with the consumer electronics manufacturers to create and promote schemes similar to the existing "HD Ready" initiative in Europe. The distance from the viewer to the screen is of course not subject to any form of control, although it seems inevitable that advice will be given on the subject, particularly where children are concerned. It may make sense to recommend both a maximum display size (beyond which objects at infinity will require correction to avoid divergence of the viewers eyes), and a maximum horizontal field of view (corresponding to a maximum screen size for a given viewing distance, beyond which the broadcast material may require the viewer to cross or diverge their eyes more than is comfortable, or exceed a comfortable ratio between the converged and accommodated distance by the viewer's eyes).

The HD era has been characterised by increases in both the largest size of television display available, and in the average size of display purchased by consumers. At present there is no reason to expect this trend to change (and indeed, NHKs plans for their "Super High-Vision" television system assume that displays will grow to fill entire walls within the next 10-20 years). Since viewing environments are not likely to grow to match, this means that stereoscopically-inclined producers of television programmes must expect both the screen size and the screen-size-to-viewing-distance ratio to increase with time. These issues could be addressed partially by the advent of high-quality view interpolation algorithms, ideally implemented in the receiver or display, and completely by the adoption of a more advanced distribution format than stereoscopy.

7 Transmission of Stereoscopic Television

The most fundamental issue for the broadcaster considering a stereoscopic service is that of what format the video takes during transmission. Display-agnosticism is an absolute requirement, at least for a broadcaster without control of the entire producation and transmission chain from camera to display. Even for broadcasters with a closed platform, adopting a transmission format tied to a particular kind of display is short-sighted: 3D display technology is likely to evolve quickly. In the short-term, display-agnosticism implies compatibility with different stereoscopic display technologies and displays within a range of likely sizes, and on reasonable assumptions about viewing conditions in contemporary domestic environments. Some of the worst problems associated with the variability of display sizes may be addressable by providing metadata alongside the video stream to allow the display to adjust the displayed screen parallax for the viewing environment, by applying horizontal images translations, for example, or scaling the displayed video down to the size of the screens for which it was intended. This would at least prevent displays larger than those targeted by the production from inducing divergence in the viewers eyes. In the medium-term, high-quality view-interpolation algorithms implemented in the display, or transmission formats based around 2D+depth representations of scenes (or variants thereof incorporating multiple viewpoints and/or other features) offer enhanced possibilities for adaptation to different display sizes, and compatibility with multi-view autostereoscopic displays to permit glasses-free viewing and a limited amount of motion parallax. At present no satisfactory general-purpose method for shooting material in this format exists; this is an area of current research in the European "3D4YOU" collaborative research project, however. In the long-term, model-based, light-field and holographic scene representations could all offer further benefits, such as a wider range of motion parallax and the elimination of the convergence/accommodation mismatch.

Existing digital television technologies such as DVB and ATSC are already more or less compatible with stereoscopic video, particularly if it is simply transmitted as two synchronous video streams, or the two streams are combined using one of the several techniques that can make them appear as a single stream to the transmission infrastructure. In the short term, such services could use existing video compression schemes, although the BBC's experiences suggest that current hardware video codecs assume that they have the freedom to drop frames arbitrarily if resynchronisation is required. This is a problem for two-stream stereoscopic video, where the video streams corresponding to the two eyes have to be kept synchronised. The behaviour of receivers when signal quality is poor must also be controlled more carefully than for two-dimensional television. For example, some digital television receivers hold the most recently decoded frame as a still if they suffer a sudden loss of signal. The effect of this occurring in one eye only of a stereoscopic broadcast would be likely to cause considerably more viewer discomfort than the monoscopic equivalent. This scenario is most likely if the video streams for the two eyes are carried in different multiplexes; something that should probably be avoided.

It would be possible to define the behaviour of receivers such that both video streams switched to a still-frame (or to a monoscopic presentation of the programme) if one stream was lost. Under conditions where the RF signal-to-noise ratio (SNR) is near to the limit of the receivers sensitivity, more complicated "error concealment" techniques can be used to deal with missing or corrupted blocks in the compressed video stream. These have been a subject of research for monoscopic video for over a decade [28], but only relatively recently has the subject of stereoscopic error-concealment strategies started to be addressed [29]. Integrating such advanced techniques into the low-margin, high-volume chipsets typically used in domestic digital television receivers is likely to be challenging in the short term, and it may be advisable to adopt a highly conservative approach, with broadcasters using higher Forward Error Correction (FEC) ratios to reduce the bit-error rate in the encoded video stream, and with receiver manufacturers encouraged to drop frames (in both eyes) or switch to a monoscopic presentation of the video in circumstances where reception is poor. A similar strategy of switching to mono if the stereo signal quality is poor is adopted by stereo FM radio receivers.

As mentioned above, a variety of schemes other than two-stream transmission could be used to broadcast stereoscopic video: anamorphically "squashing" coincident frames from the two streams into one frame in a single stream, representing adjacent frames from a 1080p25 stream as alternate fields in a 1080i25 stream, etc. The benefits of such schemes include improved compatibility with existing codecs, link and file formats. They all have disadvantages associated with them too however – typically including a lack of backwards-compatibility with existing receivers and a loss of spatial or temporal resolution. (A loss of horizontal resolution, in particular, implies a loss of resolution in the z dimension, which is compromising the very thing that the technology is trying to enable.) Some techniques are optimised for (or coincidentally matched to) specific display technologies.

Broadcasters face considerable economic and political pressure to minimise the bandwidth used by their television services, which means that simultaneous transmission schemes for stereoscopic television are likely to be superseded in the longer term: compression schemes that take advantage of the typically considerable similarity (or "redundancy") between the left- and right-eye views of a stereoscopic video stream are likely to be considered. Such "disparity-compensated" schemes are not new – the BBC patented a 3D video transmission system of this type in the 1980s [30], and the Multiview profile for MPEG-2 was defined in the 1990s [31], although it saw very little implementation, perhaps due to concerns about its efficiency [32]. Standards based on more modern codecs such as the forthcoming MPEG-MVC offer better efficiency, but even MPEG-MVC is only expected to deliver savings of around 20% compared to simultaneous transmission [33]. Understanding the theoretical limits on the bitrate savings offered by disparity compensation and the likely performance of hardware codecs over time is of the first importance to broadcasters who need to determine the timescales on which stereoscopic services could be launched.

The extent to which the bit-rate of compressed video will vary with scene content is important to designers of all kinds of video transmission system, including stereoscopic television. The corollary to the removal of redundancy via compression is that changes in the complexity of the images being compressed alter the bitrate required to compress them at a given image quality level. This is a particular problem for the hardware video encoders used for television transmission, since these must operate in real time, and cannot look ahead in the video stream to anticipate complexity changes.

This variation in required bit-rate is expected to be particularly pronounced for disparity-compensated compression of stereoscopic material, since the redundancy between the two streams will vary from complete (when a 2D image is presented) to virtually non-existent (when two different images are presented). Common scenarios where there is likely to be little redundancy between the two streams include material that has been shot with high levels of uncorrelated noise (such as sensor noise or film grain) and scenes with unusually high levels of disparity, such as an object in the extreme foreground passing in front of the cameras. In these situations, the bit-rate required to maintain quality will shoot up by a higher fraction than if the same material were transmitted without disparity compensation. If a constant bit-rate is available for that video stream, this will cause a drop in image quality. If the stream is part of a statistical multiplex "bundle", sharing a pool of bandwidth with other channels, it will cause a smaller drop in image quality on all of those channels. (By definition, an effective disparity-compensated codec will still offer improved average video quality compared to a simultaneous transmission scheme, given the constraint of fixed average transmission bandwidth for that stream, however.)

Contribution links, which bring content from "outside broadcasts" in to the broadcasters central systems for direct transmission or integration into studio-based programmes, will also require compression. The bandwidth on these links is usually considerably higher, and here adherence to published international standards is less important than the performance of the system. At present, MPEG-2 compression is commonplace (with a variety of GOP lengths); in newer hardware, more advanced codecs such as JPEG2000 and MPEG-AVC (h.264) are starting to be used. Again, the use of two standalone codecs of the kind used for 2D links would be liable to lead to synchronisation issues. This is another area in which techniques that represent two stereoscopic video streams within a single monoscopic HD stream could be used, at the cost of a significant reduction in image quality (to the point where material shot as HD should be considered Standard Definition, at least according to the BBCs criteria [24]). It would be desirable if the next generation of contribution codecs offered explicit synchronisation support for stereoscopic use, to offer broadcasters an HD-quality alternative.

The mention of quality standards for HD production leads us to consider the possibilities for "Technical Reviews" of stereoscopic content. Broadcasters perform these reviews of programmes before they are transmitted to ensure that they meets pre-defined technical criteria [34] – for example, that the audio and video signals are properly aligned, that the material is not likely to trigger photosensitive epilepsy (PSE) in viewers liable to that condition and, in the case of HD content, that the material has been originated in an approved format, and not simply 'up-converted' from a lower resolution. There are some obvious candidates for checks that should be performed as part of an equivalent process for stereoscopic content. For example, criteria could be established for rectification of the two images, for permitted limits on positive and negative parallax, for colour matching between the two video streams, and limits on the use of monoscopic material. Further research will be needed before criteria for PSE can be established for stereoscopic material. Since these technical limits restrict the artistic freedom of productions, they would have to be carefully designed in collaboration with experienced stereographers and evolved to account for both changes in the use of the stereoscopic medium and for changes in technology and display sizes.

The delivery of film and television programmes via the Internet has been an area of significant growth during the last few years, with the BBCs "iPlayer" catch-up TV service having received over 237 million requests for programmes via its website in the year following its launch in December 2007, for example [35]. A small number of manufacturers are already selling stereoscopic displays targeted at computer gamers. This implies a potential short-term audience for Internet-delivered stereoscopic video content, albeit a very small one. In the longer term, the availability of stereoscopic displays for the living room could lead to demand for stereoscopic services via IPTV. From the broadcasters perspective, the potential of the Internet for low-cost delivery to small numbers of people could facilitate experiments with stereoscopy or limited services.

8 Conclusions

The prospect of stereoscopic television gives the broadcaster much to consider. On the positive side, very little technology needs to be developed before a premium service could be launched. On the negative side, programme-makers with a track record in high-quality stereoscopy are extremely rare, and there are many artistic and engineering challenges to be overcome before stereoscopy will be both practical and affordable for productions of all genres and budgets. The biggest short-term opportunities may exist for broadcasters who can manage the broadcast chain from end to end and cover the additional costs of stereoscopy with additional revenues. Premium genres such as sport and stereoscopic film would therefore be likely to dominate early stereoscopic television services. However, broadcasters are unlikely to launch multi-genre services before consumer demand has been demonstrated, broadcast formats have been standardised and production costs are incremental compared to the costs of HD. However, as happened with HD, demand for stereoscopic production may precede services, if stereoscopic television takes off in a subset of markets.

The time for broadcasters to be experimenting with stereoscopic production and investigating how it would be displayed and how it could be transmitted is therefore well before the launch of services is forecast: HDTV again offers a precedent, with the BBCs first HD production ('The Ginger Tree' [36], co-produced with Japanese broadcaster NHK) made in 1989, preceding the launch of a BBC HD service [37] by some eighteen years.

References

[1] S Jolly, M Armstrong, and R Salmon. Three-dimensional Television: a Broadcaster's Perspective. *Proc SPIE*, 7237, 2009.

[2] R Tiltman. How 'Stereoscopic' Television is Shown, November 1928. Radio News.

[3] British Broadcasting Corporation. `http://www.bbc.co.uk/heritage/more/pdfs/1930s.pdf`.

[4] U Messerschmid, R Sand, and D Wood. Relief in Sight? - Stereoscopic television may be a realistic possibility. *Television: Journal of the Royal Television Society*, May/June 1987.

[5] L Lipton. *Foundations of the Stereoscopic Cinema.* Van Nostrand Reinhold Company, New York, 1982.

[6] R Spottiswoode, N L Spottiswoode, and C Smith. Basic Principles of the Three-Dimensional Film. *J SMPTE*, 59, 1952.

[7] R Spottiswoode and N L Spottiswoode. *The Theory of Stereoscopic Transmission & its application to the motion picture.* University of California Press, Berkeley & Los Angeles, 1953.

[8] A A Drumbreck, C W Smith, and S P Murphy. The Development and Evaluation of a Stereoscopic Television System for Remote Handling. *Image Technology*, January 1990.

[9] C Smith and A Drumbreck. 3-D TV: the practical requirements. *Television: Journal of the Royal Television Society*, Jan/Feb 1988.

[10] A A Drumbreck, C Girdwood, and M Wells. The Production and Display of a 3-D Television Programme. *ITC*, 1997.

[11] Informatics & Telematics Institute. `http://www.iti.gr/db.php/en/projects/ATTEST.html`.

[12] 3D4YOU Consortium. `http://www.3d4you.eu`.

[13] The Business Case for Digital 3D Cinema Exhibition. http://www.greenfieldscommunications.co.uk/releases/07/SDigest_3DCinema_210607.pdf, 2007.

[14] British Broadcasting Corporation. http://news.bbc.co.uk/1/hi/technology/7286852.stm.

[15] BS-11 3D schedule. http://www.bs11.jp/3d/105/.

[16] The StereoGraphics Developers' Handbook. http://www.cs.unc.edu/Research/stc/FAQs/Stereo/stereo-handbook.pdf.

[17] S Shaw. Stereoscopic 3-D - The Basics. http://www.digitalpraxis.net/stereoscopic3d-thebasics.htm.

[18] K Wittlief. Stereoscopic 3D Film and Animation - Getting It Right. *Computer Graphics Quarterly*, 41(3), 2007.

[19] Broadcast Engineering. http://broadcastengineering.com/products/3ality-digital-tackles-stereoscopic-interview-0812/.

[20] S Berezin. Stereoscopic Displays and Applications XII. *Stereo World*, 27(5):12–27, 2000.

[21] A Drumbreck. private communication.

[22] D S Cohen. James Cameron supercharges 3-D. *Variety*, 2008.

[23] KUK-film. http://www.kuk-film.de/html/nuerburg_e.html.

[24] British Broadcasting Corporation. http://www.bbc.co.uk/guidelines/dq/pdf/tv/hd_delivery_v01_09.pdf.

[25] H Postley. Comment during "Standards to enable ubiquitous 3D display" panel session at Stereoscopic Displays and Applications XX. CTO, 3ality Digital Systems.

[26] Advanced Media Workflow Association. http://www.aafassociation.org/.

[27] N Wells et al. The MXF Book, 2006.

[28] S Kaiser and K Fazel. Comparison of error concealment techniques for an MPEG-2 video decoder in terrestrial TV-broadcasting. *Signal Processing: Image Communication*, 14(6-8):655–676, 1999.

[29] C Clemens et al. A hybrid approach for error concealment in stereoscopic images. In *Proceedings of the 5th International Workshop on Image Analysis for Multimedia Interactive Services*, 2004.

[30] D H Morton. *3D Video Transmission.* June 1988. GB Patent No. GB2198310A.

[31] J-R Ohm. Stereo/multiview video encoding using the MPEG family of standards. *Proc. SPIE*, 3639:242–253, 1999.

[32] G Li and Y He. A Novel Multi-View Video Coding Scheme Based on H.264. In *Proceedings of the 2003 Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing and the Fourth Pacific Rim Conference on Multimedia*, volume 1, pages 493–497, 2003.

[33] MPEG press release. http://www.chiariglione.org/mpeg/meetings/hannover08/hannover_press.htm.

[34] British Broadcasting Corporation. http://www.bbc.co.uk/guidelines/dq/pdf/tv/tv_standards_london.pdf.

[35] BBC press release. http://www.bbc.co.uk/pressoffice/pressreleases/stories/2008/12_december/04/vision.shtml.

[36] The Internet Movie Database. http://www.imdb.com/title/tt0096595/.

[37] BBC press release. http://www.bbc.co.uk/pressoffice/pressreleases/stories/2007/11_november/30/hd.shtml.