ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/224227740

Stereoscopic 3D-TV: Visual comfort

Article *in* IEEE Transactions on Broadcasting · June 2011

DOI: 10.1109/TBC.2011.2125070 · Source: IEEE Xplore

CITATIONS
175

READS
379

5 authors, including:

Wa James Tam
Communications Research Centre Canada
92 PUBLICATIONS 2,243 CITATIONS
SEE PROFILE

Filippo Speranza
Communications Research Centre Canada
22 PUBLICATIONS 574 CITATIONS
SEE PROFILE

Koichi Shimono
Tokyo University of Marine Science and Technology
52 PUBLICATIONS 462 CITATIONS
SEE PROFILE

Hiroshi Ono
York University
169 PUBLICATIONS 2,841 CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project 2D to 3D video conversion View project

Project Depth perception View project

All content following this page was uploaded by Hiroshi Ono on 04 February 2016.

The user has requested enhancement of the downloaded file.

Stereoscopic 3D-TV: Visual Comfort

Wa James Tam, *Senior Member, IEEE*, Filippo Speranza, Sumio Yano, Koichi Shimono, and Hiroshi Ono

***Abstract*—Among the key topics of discussion and research on three-dimensional television (3D-TV), visual comfort is certainly one of the most critical. This is because it is well known that some viewers experience visual discomfort when looking at stereoscopic displays. It is important to properly address the issue of visual comfort to avoid possible delays in the deployment of 3D-TV. Here we present a concise overview of the main topics relevant to comfort in viewing stereoscopic television and survey the key factors influencing visual comfort. Potential end users of 3D-TV, content creators, program providers, broadcasters, display manufacturers and researchers will find this overview useful.**

***Index Terms*—Human factors, stereoscopic TV, three-dimensional displays, visual comfort, visual communication.**

I. Introduction

THE infrastructure for digital communication has been developing at a very rapid pace recently. On the one hand, this development has been positive for broadcasters since it has created the opportunity to deliver digital television services to multiple media platforms, which in turn permits these services to reach larger and more targeted audiences. On the other hand, the new and improved offering of television services has generated an ever-increasing competition for the attention and interest of viewers. As a result, broadcasters are being constantly challenged to innovate in order to meet customers' new expectations.

Two of the most promising new digital technologies are three-dimensional television (3D-TV) and digital cinema. In particular, the financial success of stereoscopic three-dimensional (S3D) movies has been clearly demonstrated. Recognizing the opportunity afforded by this success, the broadcasting industry has begun investigating means to deliver stereoscopic television programs and services [1], [2]. For example, the Advanced Television Systems Committee (ATSC) in North America has recently set up a planning team to examine the potential benefits and drawbacks, requirements and practical steps that are needed to deliver 3D-TV to the home. Similar investigations are also being carried out by international standard and private organizations, such as the European Broadcasting Union (EBU), the International Telecommunication Union (ITU), and the 3D@Home Consortium.

In this paper, we present a brief overview of the main topics relevant to visual comfort of stereoscopic images. The survey has been compiled with a standard television service in mind; however, the key factors influencing visual comfort apply to other services, such as internet TV and mobile TV, as well.

A. Benefits From Stereoscopic Imaging

The great interest for 3D-TV stems from the recognition that, when compared to standard two-dimensional (2D) television, this technology significantly enhances the entertainment value of television programs [3]–[11]. Clearly, the main benefit of 3D-TV is that of enhanced depth perception [4], [5], [8]. The benefits of 3D-TV, however, include more than just a greater sense of depth. Some empirical evidence suggests that stereoscopic television could also enhance the perception of sharpness [9], sense of presence [5], and naturalness [8], [10]. More importantly, surveys indicate that people would rather view S3D images than their two-dimensional counterparts [4]–[6], provided that the stereoscopic images are free from annoying artifacts and are comfortable to view [11].

B. Problems of Visual Comfort

The visual comfort of stereoscopic images has been a long-standing problem in stereoscopic research. The term visual discomfort is generally used to refer to the subjective sensation of discomfort often associated with the viewing of stereoscopic images.

Now that the demand for 3D-TV services is becoming stronger, the concerns related to the safety and health of viewing stereoscopic images have taken on a more prominent role. Some researchers have even raised the question of whether intensive watching of stereoscopic imaging could bring harm to viewers, especially to children whose visual system is still under development and, thus, probably more susceptible to external influences [12], [13]. Clearly these concerns cannot be ignored since the implementation of 3D-TV could be seriously affected if the problem of visual comfort is not satisfactorily addressed. It is therefore important to understand the underlying causes of visual discomfort so that it can be minimized or even eliminated.

The concerns are reasonable since it is well documented that some viewers experience visual discomfort when looking at stereoscopic images [14]–[20]. Already in the past, particularly in the early 50's, there have been attempts to popularize three-dimensional movies. Those attempts did not succeed because, at least in part, the limited stereoscopic technology of the time and the inclination to have stereoscopic objects pop far out from the screen often produced uncomfortable images. Therefore, it is important to properly address the problem of visual comfort so as to ensure a successful rollout of 3D-TV.

Manuscript received November 10, 2010; revised February 24, 2011; accepted February 24, 2011. Date of publication April 07, 2011; date of current version May 25, 2011.

W. J. Tam and F. Speranza are with the Communications Research Center Canada, Nepean, Ontario K2H 8S2, Canada (e-mail: james.tam@crc.gc.ca; filippo.speranza@crc.gc.ca).

S. Yano is with Japan Broadcasting Corporation, Nippon Hoso Kyokai (NHK), Japan (e-mail: syano@nhk.jp.org).

K. Shimono is with Tokyo University of Marine Science and Technology, Tokyo, Japan (e-mail: kshimono@kaiyodai.ac.jp).

H. Ono is with York University, Toronto, Ontario M3J 1P3, Canada (e-mail: hono@yorku.ca).

Digital Object Identifier 10.1109/TBC.2011.2125070

0018-9316/$26.00 © 2011 IEEE

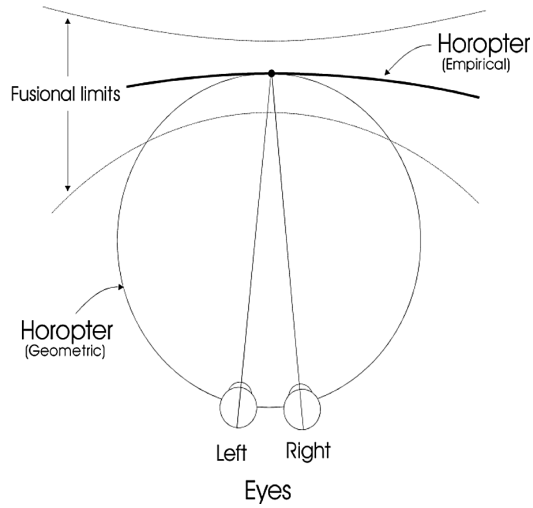

Fig. 1. Horopter and fusional limits.

This manuscript will first present some fundamentals of human stereoscopic vision and stereoscopic displays. Then it will provide a concise overview of the main issues and the key factors influencing visual comfort in the viewing of S3D video material.

II. Fundamentals of Stereoscopic Vision and Displays

A. Human Stereoscopic Vision

The ability to appreciate a third dimension using a 3D-TV display is based on the characteristics of the human visual system. Since the eyes are positioned horizontally in the head, the visual system receives two views of the visual scene, i.e., the left-eye and right-eye views, which largely overlap but differ slightly because they originate from two different perspectives. The visual system processes the information from the two images originating from the two perspectives to produce stereoscopic depth [21]–[23]. The eyes move constantly even during fixation [24]. Nonetheless, the binocular visual system is remarkably good at coordinating the movement of the two eyes [25]. As a result, from a functional point of view when we fixate binocularly a point in space, the images of that point fall, in both the left and right eyes, on the *fovea*, which is the part of the back of the eye (retina) that has the highest acuity. Thus, an object fixated binocularly is imaged on the same relative coordinates in the left-eye and right-eye views and it is perceived as a single percept, i.e., it is seen as a single object. The fixation point falls on the horopter [26], [27]. The *horopter* is a curved line or surface which contains all points that are at the same geometrical (geometrical horopter) or perceived (empirical horopter) distance of the fixation point (see Fig. 1). Like the fixated object, objects located on the horopter give rise to a single fused percept.

Points located in front of or behind the horopter are imaged at different relative positions in the left-eye and right-eye views. These differences in relative positions are termed *horizontal retinal disparities*. The magnitude of the retinal disparity of a point increases with the distance of the object from the horopter; points in front of the horopter are said to have a *negative* or *crossed disparity*, whereas object points behind it are said to have a *positive* or *uncrossed disparity*. The human visual system uses these disparities to extract the relative depth of objects in the visual scene, i.e., the position in depth of one object with respect to another object.

Objects that give rise to disparities produce disparate images on the left and right retinas. However, objects that are located within a small region in front of and behind the fixation plane still give rise to a single fused percept (see Fig. 1). The region, within which objects are fused binocularly despite having disparate images in the two eyes, is called *Panum's fusional area*. Objects located outside the Panum's area result in double vision, i.e., diplopia, but they might still be perceived in depth [28], [29]. The size of Panum's area is not fixed, rather it depends on the spatial and temporal properties of the fixation target, such as exposure duration [30], spatial resolution [31], and temporal frequency of disparity variation [32].

When the point of fixation is changed to look at a new object located at a different distance, the two eyes move simultaneously and in opposite directions so that the new object is imaged in the center of each fovea. It is in this region that fine spatial details are resolved. If the new object is closer, the eyes move inward towards each other (*convergence*), whereas if the new object is farther away the eyes move outward, away from each other (*divergence*). This process is called vergence and it is closely related to *accommodation*. The latter refers to the process by which the optical power of the eye is changed to maintain clear vision, i.e., a sharp image, of a distant object. When the eyes fixate an object, the shape of the crystalline lens in each eye is changed by the ciliary muscles so that the image of the fixated object is in focus on the back of the eye, the retina. Points located closer or farther than the accommodated point are not properly imaged on the retina and therefore subject to an increasing degree of blur. However, the visual system is tolerant of a small amount of blur, and points located within a small region around the accommodated point are perceived to be in focus. The size of this region, known as the *depth of field (DOF)*, varies inversely with pupil diameter. The depth of field has a corresponding, conjugate, region straddling the retinal plane; the region is called the *depth of focus*.

Under normal conditions, changes in accommodation of the two eyes and the process of vergence occur in an integrated fashion: changes in accommodation induce changes in vergence [33] and vice versa [34]. However, the two processes can conflict when watching stereoscopic targets, as we shall discuss later.

B. Stereoscopic Displays

3D-TV exploits the characteristics of the human binocular visual system by re-creating, albeit not in a veridical fashion, the conditions that lead to the perception of the relative depth of objects in the visual scene. Accordingly, the first requirement of stereoscopic imaging is the capture of at least two views of the same scene from two horizontally aligned cameras. The images of the objects in the scene will have different relative positions in the two views. The difference in relative position is typically called *parallax*.

When the optical axes of the cameras converge to a point in depth, the cameras are said to be in a *toed-in configuration*. The point at which the camera converges will be imaged on the same relative coordinates in the left-camera and right-camera views.

Therefore, it will have zero parallax. When displayed stereoscopically, this object point will be depicted at the screen plane. All object points located at other distances will have negative or positive parallax which will depend on the objects' distances in depth and the horizontal separation of the vantage points of the left-eye and right-eye cameras. Objects at these points will appear in front of or behind the screen plane. Toed-in configurations are easy to set up and allow an object of interest to be positioned at the screen plane but they generate *keystone distortions* in the left-camera and right-camera views. Keystone distortions transform the half-images into trapezoidal shapes such that the vertical heights no longer match between corresponding object points in the two images. These distortions might affect visual comfort.

To prevent the occurrence of keystone distortions, the cameras could be set in a *parallel configuration.* In this case, the optical axes of the cameras do not converge but rather are parallel. In the parallel configuration, all object points will have some parallax; however, images obtained with a parallel configuration could be horizontally shifted to align the images of a target object of interest to have zero parallax prior to displaying, thereby generating a parallax distribution comparable to that of the toed-in case, but without keystone distortions. Similar results could also be achieved with stereoscopic cameras employing two sensors whose offset can be varied.

To re-create the sensation of depth, the left-camera and right-camera views need to be presented separately so that the left eye sees only the left-camera view and the right eye sees only the right-camera view. This basic requirement of eye-view separation is common to all stereoscopic technologies, including the two currently being considered for 3D-TV: glasses-based active and passive systems [see invited paper by Holliman, *et al.* this issue [35]. In both of these technologies, eye separation is achieved by means of special glasses that the viewer needs to wear while viewing the stereoscopic images. The first technology uses an active time-sequential approach, in which the left and right views are temporally alternated on the display at a very high frequency. Active liquid crystal shutter glasses are used to alternatively block and unblock the visibility of the left-eye and right-eye images in synchronization with the monitor display rate. Thus, when the left view is being presented on the display, the left shutter is open but the right shutter is closed, and vice versa. As a result, each eye sees only its respective view. Slower refresh rates can introduce flicker. This however is generally not visible provided that the frequency of the change is higher than the critical flicker frequency (CFF) of the human visual system [36].

The second technology is passive because view separation is achieved by light polarization and no electronic or optical triggering signals are required. In this method the display differentially filters the left-eye and the right-eye images using light polarization; these images are then seen with corresponding polarization filters placed in front of the left eye and the right eye. Circularly polarized filters are generally preferred because they allow more head movements without affecting view separation.

Ensuring that different views are correctly presented to different eyes has proven to be quite a challenge, however. In fact, with both the passive and active technologies the eye separation is far from perfect since some of the information destined for one eye is seen by the other eyes. This leakage of information across eye views is known as *crosstalk* [37]. One of the most important requirements of a stereoscopic system is the ability to limit the amount of crosstalk. It is known that even a small amount of crosstalk can have a negative effect on picture quality [38]. However, it is less clear as to what extent crosstalk has a detrimental effect on comfort, as we shall see in Section V below.

III. What is Visual Comfort?

In the literature, the terms visual fatigue and visual comfort have been used interchangeably to describe the discomfort that might accompany the use of imaging technologies. However, Lambooij *et al.* [20] suggested a distinction between these two terms, which would make it easier to distinguish between measurement methodologies.

The term visual fatigue refers to a decrease in performance of the visual system produced by a physiological change. Therefore, visual fatigue could be assessed with physiological measures, such as changes in accommodation response [15], pupillary diameter, and eye movement characteristics [12], [39]. Visual discomfort on the other hand refers to the subjective sensation of discomfort that accompanies the physiological change. Thus, visual comfort can be measured by asking the viewer to report his/her level of perceived visual comfort.

In this review we focus on visual comfort. Accordingly, in the next section, we will briefly review some of the subjective methodologies that have been used to measure visual comfort for S3D.

IV. Measurements of Visual Comfort

Surprisingly, there are no standard methodologies for the measurement of visual comfort for stereoscopic images. For example, the International Telecommunications Union (ITU) has only one recommendation on subjective methods for stereoscopic imaging [40]; however, that recommendation only considers picture and depth quality.

In light of this deficiency, many researchers have used modified versions of the methods outlined in ITU-R Rec. BT. 500 [41] which are intended for the assessment of picture quality. Some aspects of those methods, such as presentation modes (single, double-stimulus, or continuous presentation) and sequence duration (from few seconds to several minutes duration), have been generally retained. However, in absence of common guidelines, other aspects, such as viewing conditions, criteria for material selection, and grading scales, have been inevitably different. In particular, most researchers have preferred using customized comfort scales in place of those used for picture quality assessment. For example, Tam *et al.* [42] used a double stimulus methodology and the comfort scale shown in Fig. 2(a) to evaluate the effects of reduced depth information on visual comfort. The method used was similar to the Double Stimulus Continuous Quality Scale (DSCQS) recommended by the ITU [41], and involved comparative ratings of two versions of the same stereoscopic sequence: a Test sequence whose depth information had been reduced and a Reference unprocessed version of the same sequence. Another double stimulus method was used by Kooi and Toet [43] in their investigation of the effects of binocular asymmetries on visual comfort. In their method, a

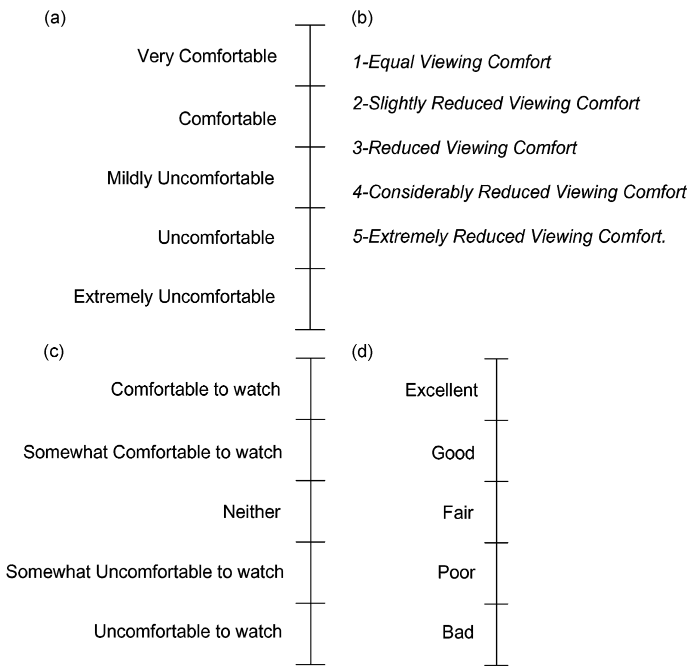

Fig. 2. Examples of scales used to measure visual comfort of stereoscopic video.

Reference target, that is, an unprocessed stimulus was always presented first and a Test target, that is, a processed stimulus, was always presented second. The viewers rated the level of comfort of the Test as compared to the comfort level of the Reference using the comparative discrete scale shown in Fig. 2(b).

Visual discomfort might vary over time, presumably increasing after prolonged exposure to the visual stimulus. Accordingly, a number of researchers have suggested the use of a continuous evaluation of visual comfort similar to the Single Stimulus Continuous Quality Evaluation (SSCQE) method described in ITU-R Rec. BT. 500 [41]. In continuous assessment the viewers are presented with a video sequence of long duration, e.g., 5, 15, or 60 minutes, and they are asked to rate the characteristic of interest, in this case the level of comfort, continuously as the image sequence is played. Therefore these methods provide a measure of "instantaneous" comfort. Nojiri *et al.* [44] and Yano *et al.* [15] used this approach to assess the effect of parallax distribution and motion on visual comfort. Their custom scale is presented in Fig. 2(c). More recently, Lambooij *et al.* [45] also used continuous assessment but opted for comfort scale similar to the ITU quality scale (see Fig. 2(d)).

The use of ad-hoc comfort scales, coupled with other methodological differences, makes it more difficult to compare across studies. Clearly, there is a need for international standard methodologies for the assessment of 3D-TV technologies, which include visual comfort measurements. International standards should also consider the complex nature of visual comfort. Indeed, the methods described above provide a uni-dimensional measure of visual comfort. However, the same level of visual comfort might be determined by different multiple sources. A number of researchers [18], [46], [47] have used questionnaires to capture the complex nature of visual comfort. These questionnaires list a series of potential symptoms or source of visual discomfort and ask the viewers to identify more precisely the source of their discomfort.

Although the subjective methods used to assess visual comfort are very similar to those used for picture quality, a fundamental difference should be noted. The assessment of visual comfort implies creating viewing conditions that might be harmful to the viewer. This has two consequences. The first is that it makes it more difficult to investigate some aspects of visual comfort, e.g., tolerance limits and long term effects. Secondly, the ethical requirements of such assessment are more stringent than those typically used in image quality assessment. In general, studies on visual comfort require more care in informing the participants of the motivations of the experiment as well as of possible negative effects resulting from exposure to the stimuli. Furthermore, it is particularly important to ensure that the participants understand that they can terminate the experiment at any time if they wish to do so.

In the next section, we will outline some of the factors that have been found to affect visual comfort and fatigue. For simplicity we will be using only the term visual comfort, or discomfort where appropriate, to indicate a change in comfort or fatigue.

V. Factors That Affect Visual Comfort/Discomfort

Despite much research, there is still no solution on how to produce stereoscopic 3D program contents that can be guaranteed to be free from visual discomfort. Nonetheless, research has identified several factors that could negatively affect visual comfort. A short description of some of these factors is presented next. For conciseness and clarity, we have grouped these factors into five categories: (a) accommodation-vergence conflict, (b) parallax distribution, (c) binocular mismatches, (d) depth inconsistencies, and (e) cognitive inconsistencies.

A. *Accommodation-Convergence Conflict*

It is generally agreed that excessive parallax causes visual discomfort [14], [17]. This is not surprising since images with larger parallaxes are more difficult to fuse.

Another possible cause of visual discomfort from excessive parallax is likely due to the accommodation and vergence conflict created by the current type of stereoscopic displays. As described in Section I, accommodation and vergence are normally yoked when viewing objects in a natural scene. However, the normal interaction between these two processes can be disrupted when viewing stereoscopic images [48]–[50]. Accommodation is directed at images of objects at the screen distance whereas vergence is directed at the perceived distances of objects.

The possible conflict between accommodation and vergence in television viewing was already investigated in the 90's by NHK (Nippon Hoso Kyokai, i.e., Japan Broadcasting Corporation). In the last two decades, this conflict has been the object of significant interest and many researchers have provided empirical evidence regarding its mechanisms [12], [15], [16], [39], [46], [48]–[53].

Under normal conditions, the distance at which the eyes accommodate and the distance at which they converge (or diverge) coincide. When we watch images on a 3D display, the viewing situation is different. The eyes accommodate to the plane of the screen so that the objects there depicted appear sharp, but they converge (or diverge) to the "perceived" location in depth of the depicted objects; an object that has negative or positive disparity will evoke a vergence response that is aimed at reducing the disparity and regain singleness of vision for that object. The larger is the disparity, the larger is the vergence response. The change in vergence in turn will elicit an accommodation response, which might cause accommodation to move away from the screen towards the point of convergence [49]. However, if accommodation moves away from the screen by an excessive amount, then the object, which is actually depicted on the screen, becomes blurred. To prevent such blurring, a corrective adjustment in accommodation becomes necessary. Therefore under stereoscopic viewing, accommodation is subject to conflicting demands whose severity depends upon the associated vergence response. These conflicting demands are often mentioned as a significant source of visual fatigue and discomfort, for the current type of stereoscopic displays.

It is generally assumed that, to minimize the accommodation-vergence conflict, the disparities in a stereoscopic image should be small enough so that the perceived depths of objects fall within a "comfort zone".

According to one approach, the accommodation-vergence conflict is reduced if the perceived depths of objects are bounded within the limits of the depth of field of the eye so that accommodation responses are minimized. Results consistent with the above analysis have been presented by Yano *et al.* [15], [16], suggesting that the depth of field can be used to define a zone of comfortable viewing. The main determinant of the depth of field of the human eye is the diameter of the pupil, which in turn varies with the level of available light. Thus the depth of field can vary substantially [53]. For the viewing conditions typical of television broadcast, researchers have assumed a depth of field, expressed in diopters (D), between ±0.2D and ±0.3D [15], [49], [55]. To illustrate, consider a viewer focusing on a TV screen located 3 meters (m) away: for a ±0.2D depth of field, the range of depth distances within which objects will appear to be in focus varies from 1.87 m to 7.5 m; for a ±0.3D depth of field, the range varies from 1.57 m to 30 m. The depth of field varies with distance: for the same ±0.2D depth of field, decreasing the viewing distance to 1.5 m will result in the range varying from 1.15 m to 2.14 m, whereas increasing it to 4.5 m will result in a range varying from 2.36 m to 45 m.

There are other approaches to characterize the limits of the zone of comfortable viewing. In one approach, these limits have been expressed in terms of maximum (positive-uncrossed and negative-crossed) retinal disparities allowed for comfortable viewing. A value of about ±1° (arc degree of visual angle) has been indicated as a possible limit by several authors [15], [18], [20], [56], in most cases on the basis of the results of empirical measurements.

Another approach uses a measure of the screen parallax, expressed as a percentage of the horizontal screen size, to specify the limits of comfortable viewing. For cinema applications, values of 1% for negative/crossed disparities and 2% for positive/uncrossed disparities (for a total value of about 3%) have been suggested [57]. Recently, some broadcasters have advocated the use of similar limits for the broadcasting

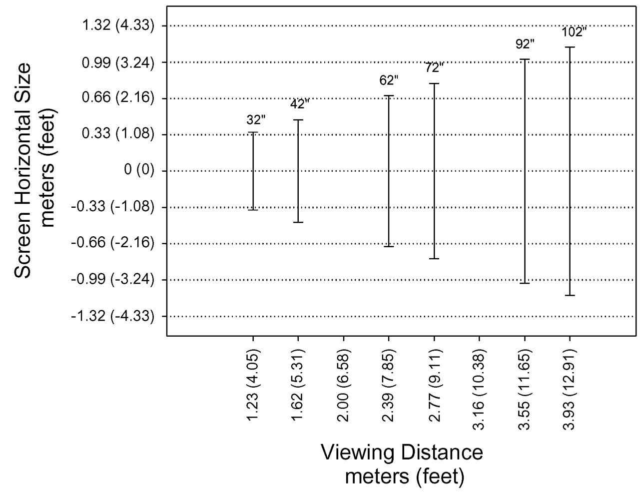

Fig. 3. Optimal viewing distance for small (32″ and 42″), medium (62″ and 72″), and large screen size (92″ and 102″) television sets.

environment [58] as well. However, it has been noted that these limits might be too small for television considering that the latter is typically characterized by smaller screen sizes than cinema; on that basis larger values, possibly as high as ±3%, have been considered [59].

For practical purposes, it is useful to compare all of these proposed limits for comfortable viewing on a common scale. For example, expressing the various limits as maximum comfortable parallax in pixels would simplify the analysis and monitoring of stereoscopic video material prior to transmission. However, identifying these comfort limits for broadcasting is not an easy task since televisions come in different sizes, are seen from different distances, and can carry different signals for image formats. Accordingly, we selected viewing conditions that simplified the measurements and allowed for easy comparison of the different approaches. For simplicity, we considered only the case of HDTV images with spatial resolutions of 1920 × 1080 pixels. In addition, we assumed that these images were viewed by an average observer from a distance that equally optimized picture quality independently of screen's size.

According to the specifications of standard organizations [60], the optimal viewing distance for the 1920 × 1080 HDTV signal is 3.1 times the picture height (3.1H). Thus, the optimal viewing distance varies with the size of the screen. It is easy to verify that for a given screen size, at the recommended viewing distance the separation between adjacent pixels (pixel pitch) subtends 1 minute of arc at the viewer's eye. The average (i.e., 20/20) visual acuity of the human eye is also approximately 1 minute of arc. Thus, at the recommended viewing distance the separation between adjacent pixels is equal to the acuity limit of the average viewer. As a result, at the recommended viewing distance the normal viewer will optimize picture quality.

We calculated the comfort limits for screens of different sizes using the optimal viewing distance as a guideline. Fig. 3 shows, in schematic form, the relation between optimal viewing distance and screen size considered here. Specifically, for each screen size we first measured the optimal viewing distance for picture quality, and then we computed for that distance the comfort limits according to the five approaches outlined above:

±0.2D depth of field; ±0.3D depth of field, ±1° of visual angle, $-1\% + 2\%$ screen parallax and ±3% screen parallax.

The comfort limits were computed as screen parallax, in pixel and mm, and disparity (visual angle subtended by the screen parallax) assuming an inter-pupillary distance of 63 mm.

Fig. 4(a) shows the limits for visual comfort in pixel values for a representative set of television screen sizes at the recommended viewing distance of 3.1H and assuming a 1920 × 1080 spatial resolution signal. Note that negative values refer to points that would appear in front of the screen plane (negative/crossed disparities) and positive values refer to points that would appear behind the screen plane (positive/uncrossed disparities). In the Figure, the solid lines represent the data for the ±0.2D depth of field; the long-dash lines represent the data for the ±0.3D depth of field; the dash-dot-dot lines represent the data for the ±1° of visual angle; the long-short dash lines represent the data for the ±3% screen parallax; and finally the short-short-short dash represent the data for the $-1\% + 2\%$ screen parallax.

The plots in Fig. 4(a) show that the three approaches: ±0.3D depth of field, ±1° of visual angle, and ±3% screen parallax, actually express nearly identical limits for visual comfort. For example, when expressed in pixel values, all three approaches correspond to approximately 60 pixels for both negative/crossed and positive/uncrossed disparities. The ±0.2D depth of field defines a medium range; when expressed in pixels this approach corresponds to about 40 pixels for both negative/crossed and positive/uncrossed disparities. Finally, the asymmetric $-1\% + 2\%$ screen parallax approach defines a much narrower range. For the negative/crossed disparities, this range corresponds to about 20 pixels, whereas for the positive/uncrossed disparities the range corresponds to about 40 pixels.

The same relationships between approaches are seen in Fig. 4(b) and (c), which present the same limits in different units. Fig. 4(b) shows the limits in terms of the screen parallax measured in mm. Note that the parallax is not constant, rather it increases with increasing screen size and viewing distance. Recall that we are considering only displays having a 1920 × 1080 spatial resolution. Therefore, given a fixed number of pixels, e.g., 60, the larger the display the larger the width in mm. Fig. 4(c) shows the limits in terms of disparity (visual angle subtended by the screen parallax). Note that the visual angle is constant for different screen sizes; this is because the larger displays were assumed to be at proportionally farther distances.

The inter-pupillary distance range varies from 45 to 80 mm in adults and can be a small as 40 mm in children down to five years old [61]. For clarity, we considered here only the average inter-pupillary distance, which as noted is about 63 mm. From Fig. 4(b), we can see that for the three approaches: ±0.3D depth of field, ±1° of visual angle, and ±3% screen parallax, the parallax can exceed the average inter-pupillary distance with larger screens, i.e., $> 92''$. Thus, these approaches might not be suitable for larger screens because parallax larger than the inter-pupillary distance, for the viewing distances considered, would cause the eyes to diverge thereby decreasing visual comfort. In this respect, the ±0.2D depth of field and the asymmetric $-1\% + 2\%$ screen parallax approaches would be more beneficial since in neither case the parallax would exceed the average inter-pupillary distance even with screens as large as the ones

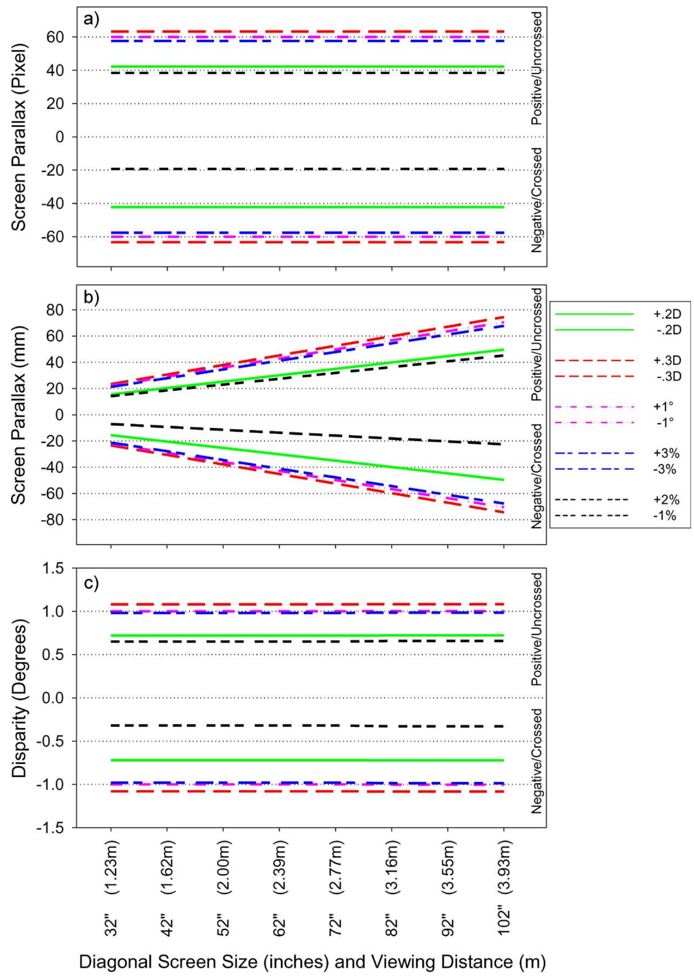

Fig. 4. Visual comfort limits in (a) pixels, (b) millimeters, and (c) disparity values.

considered here. This is because the two approaches have virtually the same comfort limits for positive/uncrossed disparities. In fact, these two approaches differ mainly with respect to the limits for negative/crossed disparities, although the difference is relatively small. Therefore, for programming directed to the general public, both approaches could be considered.

From Fig. 4(c), we can see that the disparities produced by the $-1\% + 2\%$ screen parallax proposal are rather small, particularly for negative/crossed ones, i.e., only about 0.3 degrees (20 minutes of arc). Indeed, one of the concerns with this approach that the resulting perceived depths will not be large and hence they will decrease the appeal of 3D for some viewers. Clearly, the ±0.2D depth of field approach would provide a more entertaining stereoscopic experience. However, until further research is carried out to understand the long term effects, if any, of extended exposure to stereoscopic images, the approach based on a range of $-1\%+2\%$ screen parallax might be preferable. With this range the quality of the stereoscopic experience should be sufficient; indeed some researchers have already advocated the use of smaller disparities, "microstereopsis" [62], for entertainment applications. Furthermore, in a realistic scenario the values of $-1\%+2\%$ screen parallax would represent a target limit and larger disparities would be introduced where storytelling requires it. Of course, the introduction of these larger disparities should also take into consideration the boundaries identified by experimental work, e.g., $\pm 1^\circ$ of visual angle.

The data in Fig. 4(a)–(c) show the parallax limits for visual comfort assuming an average viewer positioned at a viewing distance proportional to the screen's size. The actual viewing conditions will vary across home environments leading to differences in picture quality, visual comfort and even perceived depth [63], [64]. Since these variations cannot be predicted a priori, the broadcaster cannot adjust the stereoscopic signal to improve the quality of the service for a specific set of viewing conditions. This might be done only at the receiving end. Therefore, the limits discussed above are meant to be applied prior to transmission assuming an average home environment.

B. Parallax Distribution

Not only disparity magnitude but also the type and distribution, over space and time, of disparities seem to affect visual comfort [65].

Nojiri *et al.* [55] analysed the relationship between the parallax contained in stereoscopic still image and visual comfort. The results showed that visual comfort was highly correlated to the overall range and distribution of the parallax. These results were corroborated in a later study [44] which examined the effect of parallax distribution on visual comfort of stereoscopic HD video sequences. The results indicated that stereoscopic scenes were more comfortable to watch when the parallax distribution was such that the bottom portion of the image appeared closer and the top portion of the image appeared farther away. Furthermore, images that appeared to be mostly at the back provided greater comfort, suggesting that uncrossed disparities might more comfortable than crossed disparities. Finally, the results also indicated a decrease of visual comfort for scenes having a large parallax and large variations over time of parallax.

These latter results were consistent with the findings by Yano *et al.* [15] who examined the visual comfort of HD sequences using both subjective and objective methodologies. Viewers rated visual comfort continuously as the sequences were being played. The accommodation response was measured objectively before and after the viewing; the response decreased substantially after the viewing, although the amount of change differed across viewers. A detailed examination of the subjective rating data revealed a more significant decrease in visual comfort for scenes with large amount of parallax and large variations in the motion of objects. These results suggested that changes in disparity magnitude over time might be a major source of discomfort [16].

The negative effect of time-varying disparities on visual comfort was also observed by Speranza *et al.* [56]. These authors measured the visual comfort of stereoscopic sequences that had objects moving back and forth in depth passing through the screen plane. The objects moved at different velocities and for different ranges of magnitude of negative/crossed and positive/uncrossed disparities. The results indicated that the rate of change in disparity magnitude over time was more detrimental to visual comfort than the absolute magnitude of the crossed and uncrossed disparities.

In sum, the distribution of disparities in stereoscopic images and their change in time appear to have a significant impact on visual comfort.

C. Binocular Mismatches

With current 3D technologies there are several potential sources of binocular image mismatches. For example, the left-camera and right-camera images might be vertically offset relative to one another as a result of an improperly aligned stereoscopic camera rig; optical differences between camera lens could generate rotation and magnification errors; improper capture conditions or editing errors could create differences in luminance, color, or sharpness; capture with toed-in camera configurations would introduce trapezoidal "keystone" distortions [63]. There is much interest as to what extent these misalignment errors are potential contributors to visual discomfort when viewing stereoscopic images.

Kooi and Toet [43] examined visual comfort for several binocular image imperfections including differences in rotation, magnification, vertical offset, blur, and keystone distortions. The results indicated that distortions affecting the edges of the images, such as rotation, magnification, and keystone distortions had little effect on visual comfort, provided that the level of distortion was not high. Conversely, distortions such as blur and vertical offset reduced visual comfort even at lower levels.

Trapezoidal "keystone" distortions have received much attention. Since these distortions become larger with increased convergence, Stelmach and his collaborators [66], [67] examined the effect of convergence distance on visual comfort. They found that camera convergence distance had little effect on visual comfort except for the shortest distance examined (60 cm), which had the highest degree of keystone distortions.

The effect of absolute vertical offsets was investigated by Speranza and Wilcox [68]. In this study, viewers rated their comfort level while viewing a 3D feature film whose left and right images were vertically misaligned on a scene-by-scene basis. The feature film was presented on a large theater type screen. It was found that vertical disparities produced a marginal increase in discomfort that became only slightly more pronounced with time. These results, which suggest that the human visual system has a relatively high tolerance for vertical parallax, differ from those reported in [43]; it should be noted, however, that the latter study used a small display (home viewing conditions), whereas Speranza and Wilcox used a large display, (theater type viewing conditions). According to Allison [69], "larger displays increase the vertical vergence response and the vertical fusion range" allowing vertical disparities to "be better tolerated in large displays".

As noted, crosstalk refers to the leakage of information in the channel for one eye into the other [36]. Perceptually, crosstalk results in the so-called "ghosting" because shadows or double images are perceived near object boundaries. It is generally agreed that crosstalk is an undesirable artifact which can negatively affect the picture quality stereoscopic images [37]. It has been argued that crosstalk can also decrease visual comfort [70]. However, the available empirical evidence suggests that visual comfort is affected only when the level is crosstalk is high. Yeh and Silverstein [71] found that viewers reported moderate levels of eyestrain after conducting a stereoscopic discrimination task with a display containing crosstalk at a level of 6%–9%. Nojiri *et al.* [17] reported the results of investigations performed by NHK showing that crosstalk becomes perceptible at rather low levels (e.g., 1%) but becomes uncomfortable only at higher levels (e.g., 5%–10%). Kooi and Toet [43] examined visual comfort for still images which contained artificially introduced crosstalk levels of 5%, 15% and 25%. Comfort level decreased marginally at 5% and more substantial thereafter. More recently, Seuntiëns *et al.* [72] measured visual comfort of natural images with different crosstalk levels: 0, 5, 10, and 15%. In this case, no effect of crosstalk was found. Taken overall, these studies suggest that, at low levels, crosstalk can have a negative effect on picture quality, but it is probably not a significant contributor to visual discomfort.

D. Depth Inconsistencies

Depth inconsistencies refer to conflicting depth information resulting from errors in disparity. These errors in the depth information of the stereoscopic image might affect visual comfort.

Typically, the depth information for a stereoscopic image is embedded in the horizontal disparities between the left-eye and right-eye images. An alternative method of conveying depth information is through depth maps. A depth map is generally associated with a picture or video frame. The depth map is a matrix containing the depth of the pixels in the associated picture or video frame. The use of depth maps is based on a technique called depth-image-based rendering (DIBR) that can be used to generate new virtual camera viewpoints of a scene, given a two-dimensional (2D) image of the scene and its corresponding depth map [72]–[75]. Depth maps are a possible source of errors in depth information.

Errors in the depth map might be introduced as result of compression and/or transmission artifacts. One of the proposed methods for delivering a stereoscopic signal to the home involves a format consisting of a two-dimensional stream of images plus their corresponding depth maps [72]–[75]. This format would be very efficient for transmission and thus could make the delivery of 3D-TV signals more cost effective. Of course, the 2D+depth signal would be subject to compression and/or transmission artifacts which could introduce depth inconsistencies. However, it is not known how much such inconsistencies would affect visual comfort.

More is known about a second source of depth inconsistencies arising from depth maps: 2D-to-3D conversion. This technique computes a depth map from a standard 2D signal and uses that map to generate a 3D version of that signal. The depth maps obtained from a 2D signal inevitably contain erroneous information [77]. These errors might also result in depth inconsistencies and a decrease in comfort. Findings consistent with this expectation were recently reported by Tam *et al.* [78] in a study aimed at investigating surrogate depth maps which might contain some amount of conflicting depth information.

Temporal manipulation of depth maps might also generate depth inconsistencies. Tam *et al.* [79] conducted a study in which the depth information was updated only every n frames. It was found that dropping every other frame of depth information was well tolerated in terms of visual comfort. However, further reductions resulted in a loss of visual comfort.

E. Perceptual and Cognitive Inconsistencies

Perceptual and cognitive inconsistencies refer to conflicting information between the disparity information contained in the stereoscopic image and the depth cues that are normally experienced in the real world. These inconsistencies, even though they have no direct link to the physical and technical aspects of the stereoscopic image, might result in discomfort because they produce a cognitive inconsistency between our knowledge of the physical reality and our immediate perception [80]. To our knowledge, there has been no formal investigation that has attempted to measure the effect, if any, of these factors on visual comfort.

Edge (or window) violation is one example of such factors. Edge violation occurs when a portion of a stereoscopic object which is supposed to be in front of the screen is imaged at the edge of the screen. Since the portion of the object that is cut off by the edge of the screen may be interpreted as being occluded, this situation creates a depth conflict between the disparity cue and the occlusion cue. This conflict is assumed to lead to visual discomfort [57]. A proposed solution to the problem of edge violation is the use of a *floating window*, which consists of a virtual border perceptually located closer to the viewer than the object. The resulting perceptual outcome is consistent with the view that the virtual border is occluding the object. This method is described in [81].

The depth distortion that can occur from an off-center viewing angle might also affect visual comfort. However, it is not surprising that such distortion can create visual discomfort because even when viewing two-dimensional images, such as in a movie theatre, viewers often report headaches and visual discomfort if the viewing angle is too extreme. The degree of distortion due to off-center viewing is larger for stereoscopic images [82]. Therefore, the relevant question is whether the negative effect on comfort is also larger for 3D-TV viewing.

Patterson and Silzars [80] addressed the possible effect on visual comfort of high-level depth cue conflict. They hypothesized that "it is important that the various depth cues convey the same magnitude of depth, and thus be in registration. If they do not, then severe discomfort is the likely result." As an example, they considered a viewer watching a football game that is being broadcast on a stereoscopic display. According to the authors, the high-level cue conflict between the depth conveyed by the disparity and the distance depicted by (linear) perspective of the length of the football field "would prevent a viewer from observing the display over time without serious discomfort." Other authors have also argued that conflicting depth cues are a problem and in so much as to state that "Conflicting cues are one of the leading causes for discomfort and fatigue when viewing 3D displays" [83]. We were unable to find any empirical evidence in support of this hypothesis. Nonetheless it is plausible to expect some discomfort in presence of conflicting depth cues.

VI. Discussion and Conclusion

The rapid development of a wide spectrum of digital infrastructures for visual communication has provided an opportunity for content providers and broadcasters to carry stereoscopic

3D program material to viewers in the home. Despite the substantial progress of the last few years there are still several unresolved challenges, such as the efficient use of bandwidth to carry the 3D signals, backward compatibility with existing infrastructure, and health concerns related to the comfort of stereoscopic imaging, which could hinder the success of 3D-TV television broadcasting.

In this review, we considered one of these challenges: visual comfort. We reviewed several factors which might be at the origin of the discomfort produced by stereoscopic images, including those related to the accommodation-vergence conflict, parallax distribution, binocular mismatches, depth, and cognitive inconsistencies. We have seen that much knowledge has been accumulated over the years.

The first obvious impact of that knowledge has been in the improvement of production guidelines, e.g., [17], [57], whose key elements had been outlined in the past [84], [85]. The emphasis on production is understandable because the first step in making comfortable 3D programs is to capture them in the right way. Thus far, the movie industry has been the most active in exploring the entertainment potential of stereoscopic imaging and for that it has been rewarded with financial success by consumers. This success has renewed the interest in 3D-TV. However, the cinema and home environments differ in many aspects. Perhaps the most compelling difference is in terms of viewing conditions. The viewing conditions in movie theaters are reasonably consistent, whereas they can vary substantially across home environments: televisions come in different sizes, are watched from different distances and can display different video formats. This variety makes it difficult to control visual comfort. In other words, the same 3D program material might result in different levels of comfort in different home viewing environments.

The matter is further complicated by the fact that individual differences can result in substantially different experiences with respect to visual comfort. Some studies have found differences in terms of individual's tolerance of visual discomfort and fatigue [15], [86]–[88]. It is unclear whether these differences simply reflect normal differences in visual processes or are linked to some form of stereo-anomaly [89]–[91]. Interestingly, manufacturers are attempting to address the needs of stereo-anomalous individuals, e.g., by providing products that allow these individuals to see 2D images on 3D displays even when the latter are in stereoscopic mode.

In addition to individual differences, it is necessary to consider differences among age groups. In children, most visual functions reach adult-like levels by 6–8 years of age [92], [93], although full development might require a longer period till the teen years [94]. Similarly, visual abilities are known to decrease in older adults [95], particularly the relationship between accommodation and vergence which we have seen is very important for 3D-TV [96], [97]. Only recently researchers have begun investigating the impact of age-related differences on the visual comfort of stereoscopic video images [47].

VII. Future Direction

The dependency of visual comfort on viewing conditions, as well as the large individual and group differences suggest the need to adaptively control the depth information in stereoscopic images [98], [99].

The first step towards this control is the ability to measure objectively (i.e., using a mathematical model, the degree of perceived visual comfort in a stereoscopic stream). There has been progress made in this direction in the form of prediction systems [100] capable of measuring the presence and magnitude of spatial distortions which are known to be linked to visual comfort. This information will be necessary to devise objective tools capable of predicting visual comfort as perceived by the viewer. Some researchers have begun developing these tools [45], [101].

The capacity to measure visual comfort will create the opportunity to actually control and modify the depth information to improve comfort. In order to improve visual comfort for home viewing, it is reasonable to envision different, but complementary, options for adjusting depth information. The broadcasters might decide to edit, if required, the stereoscopic depth prior to transmission following internal or industry-wide guidelines, e.g., on the basis of the information provided in Section V above. This approach will not allow accounting for varying viewing conditions across home environments. Nonetheless, it will ensure that comfort limits are not exceeded in most common conditions.

Stereoscopic depth might be also edited at the receiver site. Depth rendering might be adapted to a specific home environment using information obtained in the environment itself by the 3D hardware, e.g., using a camera or other sensors embedded into the display, or provided by the users. Depth map estimation and depth image-based rendering (DIBR) techniques could play a significant role in this adaptation [98], [102]. These techniques could allow adapting stereoscopic movies to a specific set of viewing conditions using depth-preserving and artifact-free novel view synthesis techniques [103], [104]. Thus, the capacity to adapt depth information will eventually be part of the display system and be partially or completely under the control of the user.

Acknowledgment

The authors thank S. Daly, N. Holliman, and L. Zhang for providing suggestions for improvement on an earlier version of the manuscript.

References

[1] ITU-R, "Features of three-dimensional television video systems for broadcasting," Geneva, ITU-R Report BT. 2160, 2009.
[2] SMPTE, "Report of SMPTE task force on 3D to the home," 2009.
[3] S. Yano, "Experimental stereoscopic high-definition television," *Displays*, vol. 12, pp. 58–64, 1991.
[4] S. Yano and I. Yuyama, "Stereoscopic HDTV: Experimental system and psychological effects," *SMPTE J.*, vol. 100, pp. 14–18, 1991.
[5] W. A. IJsselsteijn, H. de Ridder, R. Hamberg, D. Bouwhuis, and J. Freeman, "Perceived depth and the feeling of presence in 3DTV," *Displays*, vol. 18, pp. 207–214, 1998.
[6] J. Freeman and S. E. Avons, "Focus group exploration of presence through advanced broadcast services," in *Proc. Human Vis. Electron. Imaging V*, 2000, vol. 3959, pp. 530–539.
[7] J. Häkkinen, T. Kawai, J. Takatalo, T. Leisti, J. Radun, A. Hirsaho, and G. Nyman, "Measuring stereoscopic image quality experience with interpretation based quality methodology," in *Proc. Image Quality Syst. Perform.*, 2008, vol. 6808, pp. 68081B-1–68081B-12.
[8] R. G. Kaptein, A. Kuijsters, M. T. M. Lambooij, W. A. IJsselsteijn, and I. Heynderickx, "Performance evaluation of 3D-TV systems," in *Proc. Electron. Imaging Image Quality Syst. Perform. V*, 2008, vol. 6808, pp. 680819-1–680819-11.

[9] A. Berthold, "The influence of blur on the perceived quality and sensation of depth of 2D and stereo images," in *ATR Human Inf. Process. Res. Lab. Tech. Rep., TR-H-232*, Kyoto, Japan, 1997.

[10] P. J. Seuntiëns, I. E. Heynderickx, W. A. IJsselsteijn, P. M. J. van den Avoort, J. Berentsen, I. J. Dalm, M. T. Lambooij, and W. Oosting, "Viewing experience and naturalness of 3D images," in *Proc. 3–D TV, Video, Display IV*, 2005, vol. 6016, pp. 601605–601607.

[11] L. Stelmach, W. J. Tam, and D. Meegan, "Perceptual basis of stereoscopic video," in *Proc. Stereoscopic Displays Virtual Reality Syst. VI*, 1999, vol. 3839, pp. 260–265.

[12] K. Ukai and P. A. Howarth, "Visual fatigue caused by viewing stereoscopic motion images: Background, theories and observations," *Displays*, vol. 29, pp. 106–116, 2008.

[13] S. Tsukuda and Y. Murai, "A case report of manifest esotropia after viewing anaglyph stereoscopic movie," (in Japanese) *Japanese Orthopt. J.*, vol. 18, pp. 69–72, 1988.

[14] M. Wopking, "Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus," *J. Soc. for Inf. Display*, vol. 3, pp. 101–103, 1995.

[15] S. Yano, S. Ide, T. Mitsuhashi, and H. Thwaites, "A study of visual fatigue and visual comfort for 3D HDTV/HDTV images," *Displays*, vol. 23, pp. 191–201, 2002.

[16] S. Yano, M. Emoto, and T. Mitsuhashi, "Two factors in visual fatigue caused by stereoscopic HDTV images," *Displays*, vol. 25, pp. 141–150, 2004.

[17] Y. Nojiri, H. Yamanoue, A. Hanazato, M. Emoto, and F. Okano, "Visual comfort/discomfort and visual fatigue caused by stereoscopic HDTV viewing," in *Proc. Stereoscopic Displays Virtual Reality Syst. XI*, 2004, vol. 5291, pp. 303–313.

[18] J. Kuze and K. Ukai, "Subjective evaluation of visual fatigue caused by motion images," *Displays*, vol. 29, pp. 159–166, 2008.

[19] M. T. M. Lambooij, W. A. IJsselsteijn, and I. Heynderickx, "Visual discomfort in stereoscopic displays: a review," in *Stereoscopic Displays Virtual Reality Syst. XIV*, 2007, vol. 6490, pp. 649001–649013.

[20] M. Lambooij, W. IJsselsteijn, M. Fortuin, and I. Heynderickx, "Visual discomfort and visual fatigue of stereoscopic displays: A review," *J. Imaging Sci. Technol.*, vol. 53, no. 3, pp. 030201–030201-14, 2009.

[21] C. Wheatstone, "Contributions to the physiology of vision.—Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision," *Philosophical Trans. Royal Soc. London* vol. 128, pp. 71–394, 1838 [Online]. Available: http://www.stereoscopy.com/library/wheatstone-paper1838.html

[22] I. P. Howard, *Seeing in Depth: Basic Mechanisms*. Toronto: Porteous Publ., 2002.

[23] R. Patterson and W. L. Martin, "Human stereopsis," *Human Factors*, vol. 34, pp. 669–692, 1992.

[24] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin*, vol. 124, no. 3, pp. 372–422, 1998.

[25] S. P. Liversedge, H. I. Blythe, and N. S. Holliman, "Binocular coordination in response to stereoscopic stimuli," in *Proc. Stereoscopic Displays Appl. XX*, 2009, vol. 7237, pp. 72370M-1–72370M-12.

[26] H. Helmholtz, "Helmholtz's treatise on physiological optics," in *Handbuch der physiologischen optik*, J. P. C. Shouthall, Ed. New York: Optical Society of America, 1925, (Original work published 1909).

[27] K. M. Schreiber, J. M. Hillis, H. R. Filippini, C. M. Schor, and M. S. Banks, "The surface of the empirical horopter," *J. Vision*, vol. 8, no. (3):7, pp. 1–20, 2008.

[28] K. N. Ogle, *Researches in binocular vision*. Philadelphia: Saunders, 1950.

[29] K. N. Ogle, "Precision and validity in stereoscopic depth perception from double images," *J. Opt. Soc. Am.*, vol. 43, pp. 906–913, 1953.

[30] G. C. S. Woo, "The effect of exposure time on the foveal size of Panum's area," *Vis. Res.*, vol. 14, no. 7, pp. 473–480, 1974.

[31] C. Schor, I. Wood, and J. Ogawa, "Binocular sensory fusion is limited by spatial resolution," *Vis. Res.*, vol. 24, pp. 661–665, 1984.

[32] C. M. Schor and C. W. Tyler, "Spatio-temporal properties of Panum's fusional area," *Vis. Res.*, vol. 21, no. 5, pp. 683–692, 1981.

[33] B. C. Jiang, "Accommodative vergence is driven by the phasic component of the accommodative controller," *Vis. Res.*, vol. 36, no. 1, pp. 97–102, 1996.

[34] M. L. Kung, T. L. Alvarez, and J. L. Semmlow, "Interaction of disparity and accommodative vergence," in *Proc. IEEE 29th Annu. Bioeng. Conf.*, 2003, pp. 29–30.

[35] N. S. Holliman, N. A. Dodgson, G. E. Favalora, and L. Pockett, "Three-dimensional displays: A review and applications analysis," *IEEE Trans. Broadcast.*, 2011, to be published.

[36] S. A. Murray and B. S. Caldwell, "Human Factors in Displays," in *The Measurement, Instrumentation, and Sensory Handbook*, J. G. Webster, Ed. Boca Raton,, FL: CRC Press, 1999.

[37] A. J. Woods, "Understanding crosstalk in stereoscopic displays," in *Proc. 3–D Syst. Appl. Conf. (3DSA)*, Tokyo, Japan, May 19–21, 2010.

[38] A. Hanazato, M. Okui, and I. Yuyama, "Subjective evaluation of cross talk disturbance in stereoscopic displays," in *Proc. 20th Int. Display Res. Conf.*, 2000, pp. 288–291.

[39] M. Emoto, Y. Nojiri, and F. Okano, "Changes in fusional vergence limit and its hysteresis after viewing stereoscopic TV," *Displays*, vol. 25, pp. 67–76, 2004.

[40] ITU-R, "Subjective assessment of stereoscopic television pictures," ITU-R Recommendation BT.1438, 2000.

[41] ITU-R, "Methodology for the subjective assessment of the quality of television pictures," ITU-R Recommendation BT.500-10, 2010.

[42] W. J. Tam, C. Vázquez, and F. Speranza, "Surrogate depth maps for stereoscopic imaging: different edges types," in *Proc. Stereoscopic Displays Virtual Reality Syst. XIV*, 2007, vol. 6490, pp. 64900C1–64900C9.

[43] F. L. Kooi and A. Toet, "Visual comfort of binocular and 3D displays," *Displays*, vol. 25, pp. 99–108, 2004.

[44] Y. Nojiri, H. Yamanoue, S. Ide, S. Yano, and F. Okana, "Parallax distribution and visual comfort on stereoscopic HDTV," in *Proc. IBC*, 2006, pp. 373–380.

[45] M. Lambooij, W. IJsselsteijn, and I. Heynderickx, "Modeling continuous assessment of visual discomfort of 3D TV," in *2nd Int. Symp. Visual Image Safety (VIMS)*, Utrecht Netherlands, Jun. 11–12, 2009.

[46] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *J. Vision*, vol. 8, no. 3, pp. 1–30, 2008.

[47] S.-N. Yang, T. Schlieski, B. Selmins, S. Cooper, R. Doherty, P. J. Corriveau, and J. E. Sheedy, "Individual differences and seating position after immersion and symptoms in stereoscopic 3D viewing," in *Proc. Stereoscopic Displays Appl. XXII*, South San Francisco, CA, USA, Jan. 24–27, 2011, vol. 7863.

[48] T. Miyashita and T. Uchida, "Cause of fatigue and its improvement in stereoscopic displays," *Proc. SID*, vol. 31, no. 3, pp. 249–254, 1990.

[49] N. Hiruma and T. Fukuda, "Accommodation response to binocular stereoscopic TV images and their viewing conditions," *SMPTE J.*, vol. 102, pp. 1137–1144, 1993.

[50] J. P. Wann, S. Ruston, and M. Mon-Williams, "Natural problems for stereoscopic depth perception in virtual environments," *Vis. Res.*, vol. 35, pp. 2731–2736, 1995.

[51] T. Inoue and H. Ohzu, "Accommodative responses to stereoscopic three-dimensional display," *Appl. Opt.*, vol. 36, pp. 4509–4515, 1997.

[52] T. Takeda, K. Hashimoto, N. Hiruma, and Y. Fukui, "Characteristics of accommodation toward apparent depth," *Vis. Res.*, vol. 39, no. 12, pp. 2087–2097, 1999.

[53] H. Mizushina, H. Ando, T. Kochiyama, and S. Masaki, "Measurements of vergence and accommodation while viewing a real 3D scene and its 2D image on a display," in *Proc. 3rd Int. Universal Comm. Symp.*, 2009, pp. 273–277.

[54] S. Marcos, E. Moreno, and R. Navarro, "The depth-of-field of the human eye from objective and subjective measurements," *Vis. Res.*, vol. 39, no. 12, pp. 2039–2049, 1999.

[55] Y. Nojiri, H. Yamanoue, A. Hanazato, and F. Okano, "Measurement of parallax distribution and its application to the analysis of visual comfort for stereoscopic HDTV," in *Proc. Stereoscopic Displays Virtual Reality Syst. X*, 2003, vol. 5006, pp. 195–205.

[56] F. Speranza, W. J. Tam, R. Renaud, and N. Hur, "Effect of disparity and motion on visual comfort of stereoscopic images," in *Proc. Stereoscopic Displays and Virtual Reality Syst. XIII*, 2006, vol. 6055, pp. 60550B-1–60550B-9.

[57] B. Mendiburu, *3D Movie Making: Stereoscopic Digital Cinema From Script to Screen*. Oxford: Focal Press (Elsevier), 2009.

[58] BSkyB, 3D technical specification for planostereocopic (3D) program content [Online]. Available: http://introducingsky3d.sky.com/a/bskyb-3d-tech-spec/

[59] Clark, "3D Production and post," 2010 [Online]. Available: http://www.itbroadcastanddigitalcinema.com/docs/3d_production_and_post.pdf, White Paper

[60] *Guidelines on Metrics to be Used When Tailoring Television Programmes to Broadcasting Applications at Various Image Quality Levels, Display Sizes and Aspect Ratios*, ITU-R Rec. BT.1845-1, ITU-R, 2010.

[61] N. A. Dodgson, "Variation and extrema of human inter-pupillary distance," in *Proc. Stereoscopic Displays and Virtual Reality Systems XI*, 2004, vol. 5291, pp. 40–75.

[62] M. Siegel and S. Nagata, "Just enough reality: Comfortable 3-D viewing via microstereopsis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, pp. 387–396, 2000.

[63] A. J. Woods, T. Docherty, and R. Koch, "Image distortions in stereoscopic video systems," in *Proc. Stereoscopic Displays Appl. IV*, 1993, vol. 1915, pp. 36–48.

[64] L. Chauvier, K. Murray, S. Parnall, R. Taylor, and J. Walker, "Does size matter? The impact of screen size on stereoscopic 3DTV," presented at the IBC Conf., Amsterdam, Netherlands, Sep. 9–14, 2010.

[65] S. Ide, H. Yamanoue, M. Okui, and F. Okano, "Parallax distribution for ease of viewing in stereoscopic HDTV," in *Proc. Stereoscopic Displays Virtual Reality Syst. IX*, 2002, vol. 4660, pp. 38–45.

[66] L. B. Stelmach, W. J. Tam, F. Speranza, R. Renaud, and T. Martin, "Improving the visual comfort of stereoscopic images," in *Proc. Stereoscopic Displays Virtual Reality Syst. X*, 2003, vol. 5006, pp. 269–282.

[67] F. Speranza, L. B. Stelmach, W. J. Tam, and R. Glabb, "Visual comfort and apparent depth in 3D systems: effects of camera convergence distance," in *Proc. 3–D TV, Video Display*, 2002, vol. 4864, pp. 146–156.

[68] F. Speranza and L. M. Wilcox, "Viewing stereoscopic images comfortably: The effects of whole-field vertical disparity," in *Proc. Stereoscopic Displays Virtual Reality Syst. IX*, 2002, vol. 4660, pp. 18–25.

[69] R. S. Allison, "Analysis of the influence of vertical disparities arising in toed-in stereoscopic cameras," *J. Imaging Sci. Technol.*, vol. 51, no. 4, pp. 317–327, 2007.

[70] S. Pastoor, "Human factors of 3D imaging: Results of recent research at Heinrich-Hertz-Institute Berlin," in *Proc. ASIA Display Conf.*, 1995.

[71] Y.-Y. Yeh and L. D. Silverstein, "Limits of fuision and depth judgment in stereoscopic color displays," *Human Factors*, vol. 32, no. 1, pp. 45–60, 1990.

[72] P. J. H. Seuntiëns, L. M. J. Meesters, and W. A. IJsselsteijn, "Perceptual attributes of crosstalk in 3D images," *Displays*, vol. 26, pp. 177–183, 2005.

[73] C. Fehn, "A 3D-TV approach using depth-image-based rendering (DIBR)," in *Proc. Visualization, Imaging, Image Process. (VIIP'03)*, 2003, pp. 482–487.

[74] C. Fehn, "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV," in *Proc. Stereoscopic Displays Virtual Reality Syst. XI*, 2004, vol. 5291, pp. 93–104.

[75] L. Zhang and W. J. Tam, "Stereoscopic image generation based on depth images for 3D TV," *IEEE Trans. Broadcast.*, vol. 51, pp. 191–199, 2005.

[76] J. Flack, P. Harman, and S. Fox, "Low bandwidth stereoscopic image encoding and transmission," in *Proc. Stereoscopic Displays Virtual Reality Syst. X*, 2003, vol. 5006, pp. 206–214.

[77] K. Klimaszewski, K. Wegner, and M. Domanski, "Distortions of synthesized views caused by compression of views and depth maps," in *Proc. 3DTV Conf.: True Vis. - Capture, Transmiss. Display 3D Video*, 2009, pp. 1–4.

[78] W. J. Tam and L. Zhang, "3D-TV content generation: 2D-to-3D conversion," in *Proc. Special Session IEEE Int. Conf. Multimedia Expo "3D-TV: Primed for Success?"*, 2006, pp. 1869–1872.

[79] W. J. Tam, F. Speranza, C. Vázquez, and L. Zhang, "Temporal subsampling of depth maps in depth image-based rendering of stereoscopic image sequences," in *Proc. Stereoscopic Displays Appl. XX*, 2009, vol. 7237, pp. 72371A-1–72371A-9.

[80] R. Patterson and A. Silzars, "Immersive stereo displays, intuitive reasoning, and cognitive engineering," *J. Soc. Inf. Display*, vol. 17, no. 5, pp. 443–448, 2009.

[81] R. K. Gunnewiek and P. Vandewalle, "How to display 3D content realistically," in *5th Int. Workshop Video Process. Quality Metrics Consum. Electron.*, Jan. 13–15, 2010.

[82] M. Banks, "Stereoscopic vision: How do we see in 3D?," presented at the NAB 2009, Las Vegas, NV, Apr. 18–23, 2009.

[83] L. Onural, A. Smolic, T. Sikora, M. R. Reha, J. Ostermann, and J. Watson, "An assessment of 3DTV technologies," in *Proc. NAB BEC*, Las Vegas, NV, Apr. 22–27, 2006, pp. 456–467.

[84] L. Lipton, *Foundations of the stereoscopic cinema*. New York: Van Nostrand Reinhold, 1982 [Online]. Available: http://3d.curtin.edu.au/library/foundation.cfm

[85] R. Spottiswoode, N. Spottiswoode, and C. Smith, "3-D photography—Basic principles of the three-dimensional film," *SMPTE J.*, vol. 59, pp. 249–286, 1952.

[86] F. L. Kooi and M. Lucassen, "Visual comfort of binocular and 3D displays," in *Proc. Human Vision Electron. Imaging VI*, 2001, vol. 4299, pp. 586–592.

[87] M. Lambooij, M. Fortuin, W. IJsselsteijn, and I. Heynderickx, "Measuring visual discomfort associated with 3D displays," in *Proc. Stereoscopic Displays Appl. XX*, 2009, vol. 7237, pp. 72370K-1–72370K-12.

[88] E. W. Jin, M. E. Miller, S. Endrikhovski, and C. D. Cerosaletti, "Creating a comfortable stereoscopic viewing experience: Effects of viewing distance and field of view on fusional range," in *Proc. Stereoscopic Displays Virtual Reality Syst. XII*, 2005, vol. 5664, pp. 10–21.

[89] W. Richards, "Stereopsis and stereoblindness," *Exp. Brain Res.*, vol. 10, no. 4, pp. 380–388, 1970.

[90] R. Patterson and R. Fox, "The effect of testing method on stereoanomaly," *Vis. Res.*, vol. 24, pp. 403–408, 1984.

[91] W. J. Tam and L. B. Stelmach, "Display duration and stereoscopic depth perception," *Canadian J. Exp. Psychol.*, vol. 52, no. 1, pp. 56–61, 1998.

[92] T. L. Lewis and D. Maurer, "Multiple sensitive periods in human visual development: Evidence from visually deprived children," *Dev. Psychobiol.*, vol. 46, pp. 163–183, 2005.

[93] S. L. Fawcett, Y.-Z. Wang, and E. E. Birch, "The critical period for susceptibility of human stereopsis," *Invest. Ophthal. Vis Sci.*, vol. 46, pp. 521–525, 2005.

[94] S. J. Leat, N. K. Yadav, and E. L. Irving, "Development of visual acuity and contrast sensitivity in children," *J. Optom.*, vol. 2, pp. 19–26, 2009.

[95] F. Schieber, "Vision and Aging," in *Handbook of the Psychology of Aging*, J. E. Birren and K. W. Schaie, Eds. : Elsevier, 2006.

[96] A. S. Bruce, D. A. Atchison, and H. Bhoola, "Accommodation-convergence relationships and age," *Invest. Ophthal. Vis. Sci.*, vol. 36, no. 2, pp. 406–413, 1995.

[97] G. Heron, W. N. Charman, and C. M. Schor, "Age changes in the interactions between the accommodation and vergence systems," *Optom. Vis. Sci.*, vol. 78, no. 10, pp. 754–762, 2001.

[98] G. Sun and N. S. Holliman, "Evaluating methods for controlling depth perception in stereoscopic cinematography," in *Proc. Stereoscopic Displays Virtual Reality Syst. XX*, 2009, vol. 7237, pp. 72370I-1–72370I-12.

[99] C. Vázquez, W. J. Tam, and F. Speranza, "3D-TV: Are two images enough? How depth maps can enhance the 3D experience," in *Proc. 16th Int. Display Workshops (IDW)*, Miyazaki, Japan, Dec. 9–11, 2009.

[100] K. Masaoka, A. Hanazato, M. Emoto, H. Yamanoue, Y. Nojiri, and F. Okano, "Spatial distortion prediction system for stereoscopic images," *J. Elec. Imaging*, vol. 15, pp. 13002-1–13002-12, 2006.

[101] S.-I. Lee, Y. J. Jung, H. Sohn, Y. M. Ro, and H. W. Park, "Visual discomfort induced by fast salient object motion in stereoscopic video," in *Stereoscopic Displays Appl. XXII*, South San Francisco, CA, USA, Jan. 24–27, 2011, vol. 7863.

[102] L. Blondé, D. Doyen, and T. Borel, "3D stereo rendering challenges and techniques," in *Conf. Inf. Sci. Syst.*, Princeton, NJ, Mar. 17–19, 2010.

[103] F. Devernay, S. Duchêne, and A. Ramos-Peon, "Adapting stereoscopic movies to the viewing conditions using depth-preserving and artifact-free novel view synthesis," in *Stereoscopic Displays Appl. XXII*, South San Francisco, CA, Jan. 24–27, 2011, vol. 7863.

[104] H. Pan, C. Yuan, and S. J. Daly, "3D video disparity adjustment for preference and prevention of discomfort," in *Stereoscopic Displays Appl. XXII*, South San Francisco, CA, Jan. 24–27, 2011, vol. 7863.

Wa James Tam (M'04–SM'07) received the Ph.D. in experimental psychology from York University in Toronto, Canada in 1990. His graduate thesis was on the effect of visual fixation on binocular eye movement latencies.

He embarked on his first project, on video coding based on viewer's eye movement patterns, at the Communications Research Center (CRC) in Ottawa, Canada in 1989. Since then he has become a senior research scientist and a project leader on 3D with the Advanced Video Systems Group at CRC. His research interests are varied but mainly centered on the theme of "applying our knowledge of the human visual system in the design and development of advanced video systems."

Dr.Tam is a senior member of IEEE and a member of SMPTE and SPIE.

Filippo Speranza received the Ph.D. in experimental psychology from the University of Toronto (Canada) in 1999.

In 2001, he joined the Communications Research Center Canada where he serves as research scientist in the Advanced Television Evaluation Laboratory (ATEL). He is currently the co-Chair of Video Quality Expert Group (VQEG), an international organization concerned with picture quality assessment. Dr. Speranza's research interests are human visual perception, 3D-TV stereoscopic imaging, and subjective picture quality assessment techniques and methods.

Sumio Yano received the B. A. and the Ph. D. degree in engineering from the University of Electro-Communications at Tokyo, Japan, in 1977 and 1993 respectively.

He joined NHK (Japan Broadcasting Corporation) in 1977, and transferred to NHK Science and Technical Research Laboratories in 1981. He was a loan employee at ATR twice: from 1992 to 1995 and from 2002 to 2004, investigating the human visual information processing. He was also a loan employee at NICT from 2007 to 2010 working on the development of three-dimensional displays. He returned to NHK Science and Technology Research Laboratories in 2010 where he conducts research on the human visual system and its application to imaging systems.

Koichi Shimono was born in Kagoshima Prefecture, Japan, on May 11th in 1955. He received the Ph.D. degree in experimental psychology from Kyushu University, Japan in 1988.

From 1984 to 1986 and 1988 to 1989, he worked at York University, Canada as a Postdoctoral Fellow and at Fukuoka University of Education as a JSPS Postdoctoral Fellow, respectively. Since 1989, he has worked at the Tokyo University of Marine Science and Technology in Tokyo, Japan where he is now a professor and member of the teaching and administrative staff. He was a visiting researcher at CRC Canada and at University of Dundee in 1996 and in 2000, respectively. His research interests are human visual and kinesthetic perception, binocular vision, visual direction and visual attention during navigation.

Hiroshi Ono was born in Kyoto, Japan and moved to the United States in 1952. He received the B.A. from Dartmouth College, New Hampshire, in 1960 and the Ph.D. from Stanford University, California, in 1965, both in psychology.

He joined the department of psychology at York University in Toronto, Ontario in 1968. He is currently a distinguished research professor emeritus at York University and a frequent visitor at ATR Human Information Sciences Laboratories in Kyoto, Japan. His research interests include the history of visual sciences, visual direction, depth and distance perception (particularly with motion parallax), binocular eye movements, and visual stability.

Prof. Ono is a member of Center for Vision Research at York University.

View publication stats