



Methods for Reducing Visual Discomfort in Stereoscopic 3D: A Review

TERZIC, K; Hansard, M

© 2016 The Author(s)

<http://dx.doi.org/10.1016/j.image.2016.08.002>

For additional information about this publication click this link.

<http://qmro.qmul.ac.uk/xmlui/handle/123456789/15098>

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk

Methods for Reducing Visual Discomfort in Stereoscopic 3D: A Review

Kasim Terzić^{a,*}, Miles Hansard^a

^a*School of Electronic Engineering and Computer Science, Queen Mary University of London*

Abstract

Visual discomfort is a significant obstacle to the wider use of stereoscopic 3D displays. Many studies have identified the most common causes of discomfort, and a rich body of literature has emerged in recent years with proposed technological and algorithmic solutions. In this paper, we present the first comprehensive review of available image processing methods for reducing discomfort in stereoscopic images and videos. This review covers improved acquisition, disparity re-mapping, adaptive blur, crosstalk cancellation and motion adaptation, as well as improvements in display technology.

Keywords: Stereoscopic 3D, Visual Discomfort, Image Processing

1. Introduction

Stereoscopic 3D is a popular form of entertainment, and is fast becoming a large industry. Stereo vision can improve performance on vision-based tasks [121], as well as audience immersion. However, many people find stereo 3D uncomfortable. Estimates of the number of people affected vary between 14% and 50%, depending on the study [144, 175]. Stereo discomfort is also known to affect viewers' emotional responses [11].

There are many symptoms of discomfort [160], which are typically associated with unnatural viewing conditions, or perceived instability of the visual world [60]. The basic causes are well understood, and their effects are quantified in a wealth of literature [90, 57, 179, 8, 185, 138, 103]. Despite this agreement on the causes, there is less consensus on how to go about improving the situation. Much research has gone into modelling and reducing these effects which is not addressed by the existing reviews. This work is spread through multiple fields, including display technology, optics, graphics, image processing, computer vision, and ophthalmology.

This paper presents the first comprehensive review of computational and technological solutions to the problem of viewer discomfort and is intended to serve as a

reference for researchers, companies and content developers interested in following this emerging field. Approximately 70% of the material presented here was published between 2011 and 2016, which indicates the recent level of activity in the field, and the need for a review at this point.

In order to understand the possible solutions, we begin with an overview of discomfort factors in this section. We follow it with a set of best practices for image and video acquisition compiled from the literature in Sec. 2. Then we discuss computational models of discomfort in Sec. 3, an essential part of any automated solution. Sec. 4 introduces algorithmic improvements intended to reduce discomfort. Sec. 5 gives a short overview of recent technological and hardware advances in display technology. We follow with a discussion in Sec. 6 and a conclusion.

1.1. Major Causes of Discomfort

There is a wealth of information in the literature about the causes of discomfort in stereo viewing [90, 57, 179, 8, 185, 103]. For completeness, we briefly introduce the main causes of discomfort here before we address models and solutions, and refer the reader to one of these reviews for a more detailed discussion of biological mechanisms underlying visual discomfort.

Crosstalk refers to the incomplete separation of images when viewing stereo 3D. Instead of one separate view for each eye, there is interference between images. Crosstalk is considered particularly annoying and it affects both depth perception and visual comfort [88]. We

*Corresponding author at: EECS, Queen Mary University of London, Mile End Road, London E1 4NS

Email addresses: k.terzic@qmul.ac.uk (Kasim Terzić), miles.hansard@qmul.ac.uk (Miles Hansard)

URL: www.eecs.qmul.ac.uk/~kasim (Kasim Terzić), www.eecs.qmul.ac.uk/~miles (Miles Hansard)

use the term “crosstalk” to refer to the physical process of interference, and “ghosting” to describe the related visual distortion, but the terms are often used interchangeably.

Inappropriate disparity is known to be a major factor in visual discomfort. Care is needed here because the term disparity is used in image processing and computer science fields to refer to on-screen disparity or parallax (in pixels), while biological community uses it to refer to retinal disparity (in degrees). For a given stereo image pair, on-screen disparity is fixed, but retinal disparity will depend on the viewing distance, viewing angle, and current gaze direction. Most of the models discussed here process images and are based on on-screen disparity so we use the term disparity to refer to the on-screen parallax. We will specify when we are talking about retinal disparity. In particular, vertical on-screen disparities (parallax) caused by misalignment of the left and right views [88], as well as views with different projections (e.g. toe-in camera configuration) [9] or sizes all contribute strongly to discomfort. Horizontal disparities are crucial for depth perception, but the human visual system struggles to fuse excessive disparities. Efficient fusion is only possible in a small region called Panum’s fusional area, close to the horopter (region of zero retinal disparity), in which the visual system perceives a single object [9]. Excessive retinal disparities are particularly difficult to fuse on the periphery of the visual field. If care is not taken, excessive parallax can lead to **window violations** (when an object appears to be in front of the screen and is cut off by the screen edge) and binocular rivalry instead of fusion [112].

With human observers, the act of fixating the gaze on a specific location in the scene is an inherently three-dimensional process, where joint eye rotation and focus are performed in order to achieve clear, single binocular vision. This is accomplished through a combination of two mechanisms. **Vergence**, is the lateral rotation of the eyes toward each other in the case of near objects (convergence) or away from each other in the case of distant objects (divergence). This process results in the object being projected onto the same area of both retinas, which facilitates binocular fusion. At the same time, the eyes adapt their focal lengths, to try bring the converged object into sharp focus, through the process of **accommodation**. In the human visual system these processes are tightly coupled (see e.g. Schor’s influential dynamical model [152]) because in natural viewing the stimuli driving the two processes are consistent with each other. But a flat stereoscopic display presents inconsistent stimuli, so the viewer tries to accommodate to one distance (the distance from the display), while at

the same time trying to converge to the on-screen disparity. This vergence-accommodation conflict is considered to be a major cause of discomfort [52].

Depth-of-Field refers to the depth range in front of the eye which appears sharp in an image. A related measure is depth-of-focus, which refers to the range of retinal defocus that can be tolerated without the perception of blur, with accommodation maintained constant [187]. It is accepted that depth-of-field can play a role in reducing discomfort [196]. In natural viewing, the eyes converge on an object of attention, bringing it into Panum’s fusional area where binocular fusion is possible. Due to the coupling between vergence and accommodation, this object is brought in sharp focus through an accommodative response. Objects behind and in front of the object of attention are blurred because of the eye’s limited depth-of-field. This helps to avoid binocular rivalry and prevents the visual system from attempting to fuse objects which, due to being far from the plane of convergence, have excessive retinal disparities. Consequently, artificial blur which simulated depth-of-field has been shown to reduce discomfort [19], and experiments show that artificial blur acts an accommodation cue and can reduce the vergence-accommodation conflict [183].

Motion can cause discomfort, but not always. In particular, fast motion in depth is known to be a major cause of discomfort [176]. This extends to sharp jumps in disparity as experienced during cuts.

Several additional unnatural effects can also contribute to discomfort. The **puppet theatre effect** is caused when the retinal disparity cues are inconsistent with the expected sizes of observed objects. For example, a person viewed on a stereoscopic display may appear to be only a metre away, but appear much smaller than a real person standing a metre away. This sensation makes the scene appear artificial. Another common complaint with stereoscopic images is the so-called **cardboard effect**, where 3D objects appear flattened in terms of depth. This effect is caused by the absence of additional depth cues such as blur along the depth gradient. Additionally, small camera baselines and limited depth resolution are known to cause this problem, which can be addressed by adaptive depth mapping techniques [165]. Finally, subjecting left and right images to different distortions has also been shown to both distort the viewers’ 3D perception [94] and cause discomfort [145]. Common examples include blur and asymmetric compression.

1.2. Display Types

Some of the discomfort factors are specifically tied to a particular display technology. Some researchers

found differences in discomfort due to e.g. using passive or active stereo glasses. Stereo displays are typically divided into displays which require additional equipment such as stereo glasses or head-mounted displays, and autostereoscopic displays which work without extra equipment.¹

Stereo glasses can be divided into passive and active glasses. Passive glasses are typically lighter and use filters designed to extract the left and right views from a combined view emitted by the main display. The most common types are **anaglyph glasses**, which use colour-based filters such as red and blue, and **polarised glasses**, which typically use circular polarisation to separate the two views. Anaglyph displays distort colours but are compatible with printing, while polarised glasses are the most popular method for stereoscopic cinema due to their low price and weight.

Active glasses are typically heavier and use time-division multiplexing, as in the case of **shutter glasses**. These are synchronised with the main display in order to alternately block the view of each eye so it does not see the image not intended for it. They are popular with some types of stereo TV systems and for computer games. High frame rates are needed to avoid annoying flicker. Additional care is needed to ensure that the views of the left and right eyes are synchronised, otherwise depth distortion can result from two eyes seeing views at different times.

Autostereoscopic displays use some kind of optical barrier or lens to ensure that each eye sees a different image. The most common types are **lenticular displays** used for larger screens, and displays with a **parallax barrier**, more popular with smaller and handheld displays. Autostereoscopic displays are attractive because they do not rely on additional viewing hardware, but they only provide a limited viewing zone. While some of them can produce different views for viewers sitting in different locations (multi-view stereo), autostereoscopic displays in general tend to suffer from high levels of system crosstalk.

Head-mounted displays use a separate display for each eye. They are typically heavier than glasses, but allow for more complete immersion and a sensation of virtual reality when combined with head tracking. They suffer from short focal distances (e.g. 37mm for Oculus

Rift) and the large field of view contributes to motion sickness with some people.

There are other types of stereo displays not addressed in this paper. Mirror and lens displays are used for some experiments reviewed here, but these are not popular for viewing today’s commercial stereo content, and are not mass-produced in the same way other displays are. There are also “true” 3D displays such as lightfields, holograms and volumetric displays which do not suffer from most of the problems discussed in this paper [51]. However, much research is still needed before they become commercially available on a large scale. In the near future, stereoscopic 3D is bound to remain the most popular way to view 3D material.

2. Improved Acquisition

Cinematographers have developed a number of rough rules for producing comfortable stereo sequences. An early set of guidelines was provided by Lipton [108] and Kitrosser [85]. Possibly the most comprehensive list of guidelines was given in Mendiburu’s book [123]. More recent papers have added to this knowledge [215, 109]. The recent review of discomfort by Urvoy et al. also provides a summary of best practices [185].

One major factor is the camera configuration. Toe-in setups were popular because they reduce the need for cropping, but this type of filming results in a mismatch between the views leading to uncomfortable eye rotation [9] and is discouraged. It is considered better to use parallel cameras with a baseline comparable to the human interocular distance, and to ensure that the cameras are horizontal and calibrated to avoid vertical parallax. Areas which are not visible to both cameras should be cropped. There are no hard and fast rules for determining the focal length because optimal focal length will depend on the camera geometry and the zone of comfort. However, geometric methods have been proposed which ensure that the camera parameters during acquisition (including the focal length) are consistent with comfortable viewing [27].

In order to prevent excessive parallax, multiple measures are given. Lambooi et al. mention the “percentage rule”, where crossed parallax (nearer than the screen) should not exceed 2–3% of the screen width and uncrossed parallax (farther) should not exceed 1–2% of screen width [90]. Shibata et al. compared this rule to their zone of comfort and recommend up to 3–4% for near parallax [163].

Modern films use considerably shorter scenes with more cuts than in the past. When a stereo scene cuts to another stereo scene, the eyes need to adjust to the new

¹Polarised glasses are popular with current LG and Sony TV models. Most Panasonic and Samsung TVs use shutter glasses, as do NVIDIA gaming products. Autostereoscopic displays are still comparatively rare, but are offered in high-end Philips 3D TVs and high-end monitors by LG and Sony. The best known HMDs are the recently released Oculus Rift and HTC Vive headsets.

depth. This can take up to 500ms, and the process is slowed down and made uncomfortable by the vergence-accommodation conflict. It is therefore recommended to use longer scenes with fewer cuts [181].

2.1. Detecting discomfort during acquisition

Heinzle et al. [49] proposed a closed-loop system capable of intuitive adjustments during acquisition in order to improve viewing comfort. They used a combination of FPGA, GPU and CPU processing to achieve real-time performance, and the resulting system makes good acquisition easier, but is not capable of post-processing videos. Sakamoto and Yakoh applied a real-time camera adjustment system capable of varying vergence for real-time operation [148]. Jung et al. used depth maps from time-of-flight cameras to detect situations leading to visual discomfort and warn the operator during acquisition [73].

3. Metrics and models

Rules of thumb described in the previous section help to reduce discomfort: a recent study showed that recent stereoscopic movies produced less discomfort than older ones [210]. But in order to develop automatic methods for improving comfort, we need computational models capable of predicting discomfort and formalising the relation between different factors and perceived discomfort. This knowledge is crucial for developing methods to automatically process stereo images in order to improve viewing experience. This section introduces recent discomfort metrics and models which can transform stereo content into discomfort scores.

3.1. Vergence-accommodation models

Early models of vergence were designed for natural viewing [153]. Vergence and accommodation in humans are coupled processes [154] which can be modelled by a dual-parallel feedback-control system [114]. Perceptual models for 2D images are not directly applicable to stereoscopic material, because they do not account for discomfort factors unique to stereo, so there is a need to develop new approaches [122]. Traditional metrics such as Zone of clear single binocular vision (ZCSBV) [41], Percival's zone of comfort [139], and Sheard's zone [159] are useful, but were designed for viewing natural scenes through lenses and prisms. Wearing lenses also leads to distortions and vergence-accommodation, but they produce a consistent modification of the visual input, which can be adapted to over time, and cues such as depth of field are still present.

On the other hand, each stereoscopic image potentially presents a different vergence-accommodation conflict, depending on image disparities and distance to the image plane.

The 'zone of comfort' model, by Shibata et al. [163] is based on a quantitative analysis of the vergence-accommodation conflict. The authors found that both Percival's and Sheard's models are too permissive for stereoscopic 3D and developed a new, stricter model capable of predicting discomfort.

Park et al. combine vergence and accommodation clues to produce a model for visual discomfort prediction [134]. They extract features which characterise vergence (disparity statistics) and accommodation (absence of de-focus blur and differential blur) and perform regression on the features to predict a discomfort value on novel images. The visual comfort model of Jung et al. combines three kinds of features: maximum absolute disparity value, maximum absolute disparity difference, and a measure of window violation [70]. The first two are directly related to the limits of comfortable viewing caused by the vergence-accommodation conflict, and the authors propose an automatic mapping method to improve comfort.

Oh et al. [132] apply a more complex quantitative model of accommodation and vergence mismatches based on responses of the fast fusional vergence mechanism. The parameters of this model are then used as features in a support vector machine trained to predict the level of discomfort.

3.2. Models based on disparity

Visual comfort for static scenes mainly depends on the screen disparity offset and range. In dynamic scenes, many factors are important, primarily the screen disparity range, lateral motion and changes in screen disparity [91]. Early discomfort models were largely based on simple disparity measures. Yano et al. used the ratio of sums of disparities near the screen to those far from the screen [202]. Nojiri et al. used minimum and maximum disparity, disparity range, and the dispersion and absolute average of disparity [130].

Kim and Sohn presented one of the first visual discomfort prediction methods. They applied first-order linear regression to horizontal and vertical disparities to produce an overall comfort score [80]. He et al. noted that discomfort varies by individual and suggested that disparity models should be personalised. They proposed calculating a Disparity Discomfort Profile separately for each viewer. Then this profile can be matched to the disparity statistics of a particular video to derive a personalised score [48].

Didyk et al. proposed a perceptual model of perceived disparity [35] based on psychophysical experiments. This model is important for disparity adjustment, because it can minimise perceived distortions in the adjusted images. The model of Park et al. is interesting because they consider both the disparity statistics and the way they are perceived by the human visual system [135]. They use coarse features derived from the statistics of binocular disparities and fine features derived by estimating the neural activity associated with the processing of horizontal disparities. The features are then analysed by support vector regression to predict a comfort score.

Jiang and Shao proposed a visual comfort based on the sparse coding paradigm [64]. They calculate disparity statistics for a static scene, and construct a dictionary based on images labelled in terms of comfort. Then they use sparse coding to represent a novel scene and derive a comfort score based on the most important dictionary elements. The same authors noted the difficulty in applying regression to mean opinion scores and suggested an approach based on preference learning, in which pairs of images are compared. They trained a classifier to pick the better looking images from a set of pairs and combined the results into a unified score [65].

Not all disparity is equally important, since not all of it attracts our attention equally. This concept was explored by Mittal et al. who combined disparity and disparity gradient with indicators of spatial activity [124] to produce the first no-reference visual quality measure for videos. Finally, Jung et al. used a model of human attention to improve their algorithm [71]. They used absolute disparity and absolute differential disparity weighted by the local salience as inputs to a regression algorithm and reported a significant improvement.

While models based on disparity have been successful, many of them make inherent assumptions about the relation of on-screen and retinal disparities. Ultimately, discomfort is caused by the projections of the world on the retina, not the physical image on the screen. These models are validated using experiments with fixed viewing distances which constrains the relation between on-screen and retinal disparities, which explains why they work. However, they do not account for different viewing configurations or important factors which affect retinal disparities such as shifting gaze. This could present an exciting new avenue of research.

3.3. Models based on motion

Several authors have explored how motion affects stereo viewing discomfort. Hoffman et al. [53] analysed motion distortion and artefacts caused by time-

multiplexing displays and found that these effects can be minimised by increasing the capture rate relative to the speed of motion and by using a single flash protocol. They also noted that time-multiplexing can distort the perceived depth of moving objects.

Human sensitivity to motion in depth follows the Weber-Fechner law [47], which states that the just-noticeable-difference between two stimuli is proportional to the magnitude of the stimuli. This insight was used by Kellnhofer et al. [75], whose disparity remapping algorithm was tuned to keep the change of disparity velocity below the just-noticeable-difference threshold. Bi and Zhou based their model on features constructed by tracking interest points using a Kanade-Lucas-Tomasi tracker and extracting salient motion depth. The features are then spatially and temporally pooled to produce a comfort measure [15].

Depth jumps. A particularly jarring type of fast motion in depth, they are common in stereoscopic video which involve cuts between scenes. The human visual system needs to adapt to such a change, which does not occur instantly. Two recent models address the temporal aspects of human response to abrupt cuts. Templin et al. measured vergence times using an eye tracker and fitted a temporal model to the data [181]. Based on their observations, larger steps in disparity lead to longer vergence adaptation times. Steps towards the screen are generally faster (because zero disparity is the special case which removes the vergence-accommodation conflict). In the model by Mu et al., response time is mainly affected by the change in disparity and increases with magnitude of the disparity. It is also affected by target disparity and target luminance contrast spatial frequency [127]. Wang et al. proposed a model for visual fatigue caused by many fast-moving objects and abrupt depth jumps [189]. They performed statistical analysis of spatial characteristics, temporal characteristics and scene movement characteristics, extracted from salient regions, then applied linear regression to the resulting features to predict visual fatigue.

3.4. Crosstalk models

In order to eliminate crosstalk, it must first be identified. Huang [58] distinguishes between system crosstalk (related to the device) and viewing crosstalk (related to the content). System crosstalk depends on the display itself and can therefore, at least in theory, be removed through calibration. Weissman and Woods [194] presented a simple way to measure crosstalk by viewing a test chart. An early crosstalk model is given by Konrad et al. [87], who use an additive model which

combines the intended image and a crosstalk term calculated from the intended image and the other view. This model can be extended to colour images by applying it to each colour channel in RGB space separately, but the results do not always agree with human perception. Kang et al. [74] noted that a model based on lightness difference works best when the intended image is black, but when the intended image is not black, colour difference (in the CIELAB space) works better. Zhang and Shen also presented a method for predicting crosstalk in colour images without measuring it. They combine the disparity map, colour difference map and colour contrast map from original stereo images to drive their model [212].

Perceptual models. In addition to measuring or predicting the amount of crosstalk, it is important to model how it is perceived by human viewers. Seuntiens et al. presented an early analysis of crosstalk perception [156] which found that crosstalk is more visible at larger camera separations. Based on this insight, Xing et al. built a perceptual model which combined crosstalk level, camera baseline, and scene content to predict user scores in a Quality Assessment scenario [199].

An essential problem for reducing content-based crosstalk is determining the visibility threshold for crosstalk. Wang et al. examined how contrast and binocular disparity influence crosstalk perception and presented an analytical formula for predicting the visibility and acceptability threshold for crosstalk [191]. A recent study found that crosstalk metrics based on the Weber-Fechner law correlate well with human perception [198]. In contrast, Shestak et al. performed threshold detection based on a nonlinear Barten's model [161].

Crosstalk models are closely coupled with crosstalk cancellation. Several crosstalk cancellation methods are discussed in more detail in Section 4.

3.5. Other factors

One of the earliest visual comfort models was based on the discrepancy between the left and right views caused by distortions such as blur, vertical on-screen disparity or image compression [145]. This model was good at predicting discomfort caused by these less explored factors but did not address the main causes of discomfort such as excessive disparity and motion in depth.

Sohn et al. demonstrated that object thickness plays an important role in discomfort. They proposed a model which combined disparity magnitude to improve prediction performance [169]. Their model was later extended to using object size and relative disparity information

[170]. Chen et al. predicted the scene quality based on a stereoacuity model [25]. They demonstrated that different processing is needed for foreground-dominant and background-dominant images and proposed an automatic method for determining optimal disparity shifts.

Finally, it is important to note that visual discomfort is a dynamic phenomenon, and that it changes with time. Kim et al. [84] introduced a temporal visual discomfort model. They model neural activity by a second-order differential equation and perform an analysis in the Laplace domain to examine its stability.

3.6. Models combining multiple cues

In recent years, most models have been built on combinations on several cues in order to improve prediction performance. However, it is not at all obvious how the cues should be combined. An early model by Lambooij et al. [90] used linear regression to combine disparity range, lateral motion and disparity change into a unified score. Choi et al. [32] considered a large number of possible predictors such as spatial and temporal complexity, motion, brightness, crosstalk and depth gradient, and then performed principal component analysis to select significant factors. Then they used multiple regression to predict discomfort factors, followed by a weighted linear combination. In a similar vein, Lee et al. [95] introduced a 3D visual activity framework which captures statistics of 3D scenes, including colour, texture, motion and disparity. They showed that these features can be used as a predictor for visual discomfort.

In one experiment, Minkowski summation combined with a high exponent and max-combination yielded the best accuracy in predicting the overall level of visual discomfort. This suggests that discomfort is dominated by the most significant discomfort factor, i.e., the winner-take-all mechanism [97]. Finally, Chen et al. suggest that visual discomfort is not a linear function of discomfort factors and propose applying the Weber-Fechner law [29]. Some recent models apply machine learning methods to this problem, by careful feature selection followed by some form of regression (usually Support Vector Regression – SVR) designed to relate the features to subjective discomfort scores, as in [31].

Models based on saliency. Following a somewhat different approach, Iatsun et al. [61] developed a model for predicting discomfort based on eye tracking data: fixations, blinks and focus. They found that discomfort strongly correlated with spatial saliency, motion intensity and disparity range. Cho and Kang [31] used object salience derived from a region-based algorithms. Then

they extracted features including disparity, motion, contrast, spatial complexity of salient objects and brightness and binocular asymmetries degree between left and right image to construct their model. The features were fed into an SVR to predict discomfort.

Several models have used saliency maps to improve discomfort prediction, but it is also true that discomfort drives attention. Jiang et al. [66] proposed a 3D saliency model which explicitly accounts for visual discomfort. They combined colour saliency, texture saliency and spatial compactness with global disparity contrast to train a comfort prediction function which classifies scenes into High-Comfortable (HCVS) and Low-Comfortable Visual Scenes (LCVS) and used this information to generate a saliency map based on viewing comfort. Kim et al. [81] add a predicted discomfort score to other saliency attributes such as motion, disparity and texture, in order to create a refined 3D saliency map.

3.7. Quality Assessment

An overview of comfort models would not be complete without mentioning the available quality assessment algorithms. These were traditionally concerned with artefacts caused by coding or compression, but are typically mapped to subjective human evaluation, and this is strongly correlated to 3D factors discussed in the rest of this paper. That said, QA methods are mapped to human opinion of image *quality* and not *comfort*, and these are sometimes conflicting criteria. Modern quality assessment methods have been adapted to stereoscopic content and, while they do not model discomfort caused by 3D effects explicitly, discomfort measures are typically integrated into a complete QoE score and are therefore relevant to this discussion. Due to the number of recent assessment methods, there is increasing need to standardise QoE across labs [10]. For an overview of quality assessment methods for stereo 3D, we refer to [195]. In this section, we focus on methods that explicitly incorporate visual discomfort.

It has been shown that the inclusion of disparity error in QA models improves their correlation with human responses [5], as does modelling binocular rivalry [26]. Yun et al. produced a no-reference stereo assessment method without using external depth maps [205]. Ryu et al. incorporated stereo asymmetry [147]. Both were shown to improve correlation with human scores. Park et al. defined a set of universally relevant geometric stereo features, and built a regression model that effectively captured the relationship between stereo features and the quality of stereo images. They report that their model performs on a par with the average human [137].

Ha and Kim introduced a metric based on temporal variance, disparity variation in intra-frames, disparity variation in inter-frames and disparity distribution of frame boundary areas [46]. Shao et al. classify regions into non-corresponding, regions of binocular rivalry, and regions of binocular fusion [158] and evaluate them independently before calculating a combined score.

Much research has gone into perceptual evaluation of compressed video. It is known that depth discrimination in humans is less strong than along the spatial dimensions, but excessive depth compression can result in unnatural distortion. Pajak et al. introduced a perceptual model for depth compression [133].

3.8. Datasets

We close this section with a set of widely available stereo datasets commonly used to assess visual comfort. With the increased interest in the causes of discomfort, researchers have moved from synthetic depth-corrugations to large databases of natural images, mirroring recent trends in Computer Vision and Machine Learning.

A number of quality assessment stereo datasets are available [213], which can be used as a basis for psychophysical experiments and evaluating novel algorithms. The dataset by Corrigan et al. collects typical broadcasting material combined with typical distortions [33], but these datasets do not include opinion scores, so each lab must perform a separate psychophysical experiment, which makes the results difficult to compare.

The LIVE 3D IQA database by Moorthy et al. combines stereoscopic pairs with depth maps and subjective opinion scores [126, 125]. The EPFL datasets of stereo images [44] and videos [43] by Goldmann et al. also include subjective quality scores. The recent multimodal QoE dataset by Perrin et al. [140] includes subjective scores as well as EEG and ECG readings of the viewers. Depth maps are error prone. Kondermann et al. presented a 3D video dataset where the provided depth maps have error bars modelling their uncertainty [86].

In terms of addressing discomfort specifically, the recent IEEE SA standard 3333.1.1 [2] defines a database of images and videos for evaluating discomfort based on psychophysical experiments [3], and several recent papers based their evaluation on these images. The KAIST dataset [1] comprises 120 stereo images with associated discomfort levels, expressed as Mean Opinion Scores obtained from 17 test subjects. The SCCH dataset [4] comprises 21 stereoscopic videos with subjective scores for discomfort, depth level and image quality, also expressed as Mean Opinion Scores from 17 test subjects.

These databases are likely to become the standard for discomfort measurement in the near future.

4. Algorithmic Improvements

Some causes of discomfort can be fixed during acquisition by following best practices outlined in Section 2. But this is not always sufficient for several reasons. Firstly, many modern movies are produced in 2D and 3D at the same time, and directors may not wish to compromise their artistic vision by adhering to a much more restrictive set of guidelines specific to stereoscopic video. Furthermore, there is much material that was filmed before such guidelines were established, and even with modern films, discomfort is still an issue. Finally, comfort is highly dependent on the display device, and stereo content needs to be comfortable to view on a variety of devices, which pose different constraints. This makes the retargeting of stereoscopic videos a very important area of research, and many of the advances in terms of view synthesis, remapping and registration have made their way into commercial products for movie development [40].

On mobile devices, viewers may prefer depths either behind or in front of the screen-plane [164], depending on their degree of exophoria (the tendency of the eyes to deviate outward with respect to focal distance). It follows that depth remapping may need to be user-specific, and to operate in real-time.

4.1. Disparity range mapping for comfort

Yan et al. [201] define apparent depth Z from stereoscopic video as

$$Z = \frac{ab}{b-d}, \quad (1)$$

where a is the distance from the screen, b is the interaxial distance, and d is the on-screen disparity. Disparity mapping can then be seen as a process which transforms the apparent depth Z with range $[Z_{\min}, Z_{\max}]$ to a target depth \hat{Z} with range $[\hat{Z}_{\min}, \hat{Z}_{\max}]$:

$$\hat{Z}(\mathbf{x}) = W(Z(\mathbf{x}) - Z_{\min}) + \hat{Z}_{\min}, \quad (2)$$

where W is a mapping function. Disparity mapping algorithms mostly differ in i) the definition of W : linear or non-linear, local or global; and ii) the process by which the two stereo images are transformed to obtain the desired disparity: warping, shifting, and so on.

Early dynamic depth mapping algorithms were limited to 3D rendering, where scene geometry is known. Probably the first dynamic mapping algorithm was proposed by Ware et al. [193]. They noticed that subjects

who were allowed to manually adjust the amount of disparity tended to prefer a certain range of depths and proposed a method for modifying 3D scenes rendered by a computer.

It has been suggested that camera separation equal to the interocular distance in humans is preferable because it most closely resembles natural scenes. However, special care is needed for panoramic images, because different directions need different camera baselines to ensure comfortable viewing [141]. Jones et al. introduced a geometrical model which allows a camera operator to choose optimal camera parameters including the baseline width in order to produce disparities within the acceptable range for 3D rendered scenes, and also for real cameras. Their work also allows for real-time processing to account for free head movement of the observer [68]. Sun and Holliman also presented an algorithm to automatically scale the disparity range for 3D renderings. They use the Z-buffer from OpenGL to dynamically control depth by varying inter-camera distance [177]. Li et al. reconstructed the scene geometry from multiple views and devised a method for repositioning the camera and scaling depth in a way that reduces viewing discomfort [101]. It built strongly on the multi-view synthesis presented in the previous section.

Where full scene geometry is not known, it is useful to have a separate depth map. Kim et al. extended the disparity scaling principle multi-view stereo. They use external depth maps to help disparity calculation and use inpainting to cover the holes resulting from shifting pixels [83]. A depth map can be automatically generated from stereoscopic images in order to manually manipulate depth and generate arbitrary stereo pairs [188]. Wang et al. improved depth mapping for Depth-Image-Based Rendering, a common stereo transmission format [190].

Cho et al. introduce a saliency measure in order to reduce visible warping artefacts [30]. They introduce an energy function combining geometric distortion and alignment consistency. They also include a term to limit maximum disparity in order to improve viewing comfort. The warping-based approach by Lin et al. also incorporates saliency maps and segmentation [106]. It also adds a cropping stage to detect and remove window violations which are known to be very uncomfortable. Jung and Ko adjusted disparities of neighbouring objects to be above the just-noticeable-difference threshold [69]. The result was more pleasing to view because there was more perceived depth, but they did not address excessive disparities or viewing comfort.

Disparity shifting. Most image warping techniques used for depth mapping produce visible artefacts. A fast alternative is to shift disparities so that they are centered around the display plane. Qi and Ho advocate very fast stereo retargeting by using ‘shift maps’ to move the zero-disparity plane [143], which was found to improve viewing comfort. Shao et al. use spatial frequency, disparity energy and visual attention to predict the optimal zero-disparity plane. Then they shift the images so that this plane coincides with the display plane [157] and then apply depth scaling to avoid excessive disparities, which leads to less visual distortion. In a similar vein, Jung et al. (whose visual comfort metric was discussed in Sec. 3) first shift then remap [70]. They also include a cropping stage to avoid window violations, as in [106]. Kim et al. measured the time required for fusion under different viewing conditions and used it as a comfort predictor. They used SURF keypoint correspondences extracted from a novel stereo video to estimate disparity distributions and applied face detection to estimate the viewing distance. Then they shifted the images in real time to reduce viewing discomfort in real time [79].

As discussed earlier, the depth range of stereoscopic videos needs to be within a relatively narrow range to ensure visual comfort. But aggressive depth scaling can produce visible distortions which detract from perceived visual quality. Two proposed solutions are non-linear scaling and local depth adjustment. Sohn et al. proposed local disparity remapping by splitting the depth map into a coarse and detailed map, and applying a model of stereoacuity to process objects locally without introducing distortion [174].

Nonlinear scaling. Nonlinear depth scaling methods for stereo retargeting have existed for a long time [39]. They have only recently been applied to reducing viewing discomfort. Lang et al. proposed a nonlinear mapping function based on perceptual insights [92]. It is a method based on sparse correspondences and image warping. They defined a set of disparity mapping operators which can be combined to provide a non-linear and locally adaptive depth mapping which attempts to match the target disparity range while maximising viewing comfort. A similar, but simpler nonlinear mapping was proposed by Wu et al. [197] and patented in 2012 [24]. They used a squashing function which is linear for average image disparities and strongly non-linear for extreme disparities and reported an improvement in viewing comfort. Sohn et al. proposed a method which combines global linear scaling with local nonlinear scaling for problematic regions [172]. The scene is first linearly compressed into the desired disparity

range, then depth planes which strongly contribute to discomfort (e.g. objects with strong disparity gradients) are locally processed using a nonlinear operator. Both stages are iterative and continue until some target comfort level is achieved. Oh et al. proposed a very similar method. They predicted visual fatigue based on spatial frequency, disparity magnitude and disparity motion, and then applied non-linear remapping to reduce fatigue [131]. The main difference from [172] is in the perceptual models and mapping functions used, and that they explicitly enforce temporal coherence in moving scenes.

Personalised mapping. Since humans differ in terms of depth perception and visual discomfort, several user-tailored dynamic mapping approaches have been proposed. Mangiat and Gibson introduced automatic disparity remapping for 3D video telephony on handheld devices [118]. Their main insight is that the object nearest to the camera (in this case, usually the head of the person being called) should be placed on the display plane. Bernhard et al. [14] proposed fast disparity adjustment based on gaze tracking which is personalised for each user (see also [67]). They only adjust disparities at extremities or outside of the user’s measured comfort zone. The retargeting method by Masia et al. is adapted to a particular display rather than a particular user. They address the problem common with autostereoscopic displays whose depth resolution is very limited, leading to strong blurring for disparities outside of a narrow range [119].

4.2. Artificial Depth of Field

A significant amount of work has gone into dealing with depth of field in stereoscopic viewing. Early experiments with simulated depth of field by Wöpping showed that blurring of non-fixated areas increases viewing comfort [196]. Further experiments with mirror displays by Blohm et al. showed that viewers preferred rendered scenes where only a small sub-volume containing the objects of interest is presented in full resolution, which they termed Depth of Interest (DoI) [17]. Artificial depth of field algorithms can be seen as generalized blurring operations $I(\mathbf{x}) = K_{\mathbf{x}} \odot I$, on the image I . Here the kernel $K_{\mathbf{x}}$ at position \mathbf{x} depends on the actual distance-map Z , and on the focus distance Z_f , hence:

$$I(\mathbf{x}) = K_{\mathbf{x}}(Z(\mathbf{x}), Z_f) \odot I. \quad (3)$$

The proposed approaches differ in determining the desired focus distance (e.g. through eye tracking or salience models) and the dependence of the operation on \mathbf{x} .

Eye tracking. Artificial blur was combined with eye tracking by Talmi and Liu [178], resulting in an autostereoscopic display capable of simulating depth of field in real time. With increasing availability of inexpensive eye trackers, a number of researchers have built similar systems since then. Hillaire et al. applied this idea to 3D games and found that it improved immersion [50]. Vinnikov and Allison [186] and Duchowski et al. [38] recently presented novel displays capable of simulating DoF based on eye tracking data. The latter system, called the Gaze Contingent Display (GCD), was designed specifically for improving comfort, but their evaluation showed no significant improvement in comfort. Moreover, like Vinnikov, they found that viewers prefer sharp images to artificial DoF.

Two main reasons are thought to be imprecision of the eye trackers and the small, but perceptible delay between eye movement and the response of the system. Another reason might be that recent systems [186, 38] keep viewers' heads fixed, leading to discomfort. Improvements in eye tracking performance and accuracy will hopefully resolve these questions in future experiments.

Selective blurring. Eye tracking and selecting filtering in real time has the disadvantage that it only works for a single viewer. Alternative approaches have attempted to apply selective blur in a way that is viewer-independent and can be used for TVs and cinemas. It is known that depth of field is especially helpful with high frequency content associated with large disparities, which leads to binocular rivalry and diplopia. Leroy et al. described an algorithm which selectively blurs areas with high disparities and showed that it improves viewing comfort, but that the output is less aesthetically pleasing than sharp images [98]. It was shown in a similar experiment that active blurring improves fusion and comfort on both a stereo display and a see-through HMD [18]. An alternative was explored by Jung et al. [72] In their approach, they use a salience operator to selectively blur areas deemed less important. This is essentially averaging over a large number of fixations and viewers and blurring areas less likely to be fixated.

A simpler model was applied for interactive games with HMDs by Carnegie and Rhee [19]. They applied dynamic DoF filtering by focussing on the centre of the screen, arguing that i) this is where people will focus most of the time and ii) focus acts as an attentional cue, so people will be encouraged to look at the centre of the screen more often. Additionally, they simulated the refocussing delay to make focus transitions less jarring. They found that there was a reduction in sickness and

discomfort for many, but not all participants.

4.3. Motion

Although motion in stereo films is one of the major causes of discomfort, comparatively little work has gone into algorithmic solutions. Motion adds a further cause of discomfort which might have a complicated dependence on other causes of discomfort (e.g. excessive disparity). In video, disparity maps can be seen as functions of position and time: $Z(\mathbf{x}, t)$. In addition to detecting excessive disparities in each frame (which depend on the position \mathbf{x}), it is necessary to remove sharp changes in disparity between consecutive frames at time points t and $t + 1$. Typically, algorithms for reducing discomfort will define a cost function $c(Z(\mathbf{x}, t), Z(\mathbf{x}, t + 1))$ such that large changes in disparity are penalised. Reducing discomfort then amounts to finding a set of disparity maps $\hat{Z}(\mathbf{x}, t)$ which minimise this cost function.

Sohn et al. addressed the problem of fast motion in depth [171]. Their algorithm detects fast changes of disparity in visually salient regions and adjusts using local disparity remapping. The focus on visually salient regions partially helps to preserve the natural feeling of the scene for human observers, but disparity mapping is traditionally frame-based, which can result in visible distortion of smooth motion in depth. Remapping stereoscopic video needs to take motion consistency into account, such as using optical flow between successive frames, as done by Lang et al. [92]. Kellnhofer et al. proposed warping model applied to a space-time cube, which takes account of longer movements in depth and results in more natural motion [75]. The natural motion is a result of performing a global optimisation over the entire scene rather than frame by frame. The previously discussed disparity mapping approach by Oh et al. also explicitly takes motion into account [131]. They use velocity of motion in depth as one of the factors for predicting visual fatigue, which drives the disparity mapping process, but note that their method produces distortions for objects which move in depth.

A special case of fast movement in depth occurs when one scene abruptly cuts into another. A simple solution is a fade to zero disparity [89]. Delis et al. proposed an algorithm for automatically detecting depth jumps based on average positive and negative disparities [34]. Another special case occurs when motion in depth leads to excessive disparities and finally window violations. Nazzer et al. presented an algorithm for automatically detecting such events and issuing a warning [129].

Kellnhofer et al. addressed two causes of false motion: the Pulfrich effect common with anaglyph glasses

and false motion caused by time-multiplexing displays [76].

4.4. Crosstalk Reduction

Algorithmic solutions for crosstalk have been around at least since Lipscomb and Wooten [107]. Since both left and right views are known during playback, it is typically possible to model the amount of “leakage” from one view to the other during a calibration step, and to subtract a scaled version of the left view from the right view (and vice-versa) during playback or postprocessing, a process called “crosstalk cancellation”. Konrad et al. [87] give a general form of the process:

$$f'(\mathbf{x}) = f(\mathbf{x}) + \phi(f(\mathbf{x}), g(\mathbf{x})), \quad (4)$$

where f is the intended image (e.g. left view), g is the interfering image (e.g. right view), and ϕ is a crosstalk function. The majority of algorithmic approaches today use a variation of this process, and they differ by the exact definition of the crosstalk function ϕ .

Since it is based on image subtraction, cancellation reduces the overall brightness of the scene, and can result in washed-out colours so many algorithms attempt to counteract this. It also fails in so-called “uncorrectable regions”, where the contribution of crosstalk is larger than the contribution of the correct image. This is typically handled through intensity adjustment, the simplest form being global mapping [87], but this type of mapping is known to reduce contrast. More recent local contrast reduction is capable of reducing crosstalk while preserving dynamic contrast in a scene [36].

Much early work on crosstalk reduction assumed greyscale images. When these algorithms are applied to the R, G, and B channels separately, it results in distorted colours [74]. An alternative is to scale the luma channel in the YCbCr colour space to reduce crosstalk without distorting the colours [37]. Zeng et al. showed that crosstalk can be completely removed by using linear programming in the YCbCr space [208].

Unlike disparity, crosstalk is highly dependent on the particular type of display technology used, and different solutions were proposed for different displays (some hardware improvements are discussed in Section 5). Early crosstalk cancellation in of Konrad et al. was applied to time-sequential displays [87] but work on time-sequential displays is usually combined with hardware advances. Anaglyph stereo systems traditionally suffer from strong crosstalk. This can be reduced through heuristic thresholding [149] or blurring [62]. If the spectral distribution of the display device and the transmission functions of the anaglyph filters are known,

crosstalk can be calculated and removed [120], but this information is not always readily available. Sanftmann and Weisskopf presented a quick calibration method for anaglyph stereo with five parameters based on a perceptual luminance model [150]. Zeng et al. developed a similar model for circularly polarised LCDs [207, 209]. Like their earlier models, they used linear programming to derive optimal images for crosstalk correction.

Crosstalk is exacerbated by large disparities, so disparity mapping can also be applied to reduce crosstalk. If disparity is adjusted so uncorrectable regions are aligned, then ghosting can be eliminated for that plane [62], but this is not useful for the general case because it might increase ghosting at other depths. The solution by Sohn et al. combined disparity mapping and crosstalk cancellation [171, 173]. They first detect the areas in the image where crosstalk is hard to cancel and reduce disparity in these regions and then proceed with cancellation.

Lenticular lenses are commonly used to construct multi-view autostereoscopic displays. Such displays suffer from ghosting caused by leakage from neighbouring pixels and present the additional difficulty of having to model the complex interplay between many different views. Chang et al. generalise the idea of view subtraction to multi-view 3D by constructing a crosstalk matrix [22]. More recently, Wang and Hou presented a crosstalk calibration and removal system for lenticular displays by formulating it as a box-constrained integer least squares problem [192]. Li et al. model crosstalk between vertical neighboring subpixels by a shift-invariant low-pass filter, and propose a filtering method in the frequency domain to reduce ghosting on lenticular 3D displays [102]. Finally, Zhou et al. presented a unified method for intrinsic (leakage from neighbouring pixels) and extrinsic (due faulty manufacturing) crosstalk for slanted lenticular displays [214].

Most crosstalk cancellation methods process pairs of images, ignoring temporal aspects. In practice, this leads to jitter with fast-moving objects. Smit et al. solve this through a dynamic model that takes movement into account [167]. Their non-uniform model also accounts for different amounts of crosstalk in different parts of the image. Hong [55] also noted that users wearing shutter glasses do not experience the same amount of crosstalk when viewing from different positions, and proposed an algorithm to correct for this effect.

With the rising popularity of handheld displays, active methods are becoming important. Some algorithms are capable of running in real time, but latency could be a problem for interactive scenarios. Crosstalk cancellation was successfully implemented on an FPGA

board [78]. Increased computing power has made active crosstalk cancellation based on viewing angle for handheld devices possible [23].

Finally, several recent patents have addressed crosstalk. The first method constructs a third “margin” image and uses it for active crosstalk cancellation [136]. Later methods remove crosstalk either by directly modifying one of the views [77], or both of them [204].

4.5. Stereo Retargeting

With the proliferation of different viewing devices, same content may be viewed on displays as large as a cinema screen or as small as a mobile phone. This gave rise to the vast field of content retargeting, some of which has been extended to stereo content. Most of these are primarily concerned with mapping to a specific device without regard for viewing comfort. However, since they aim at producing a pleasing target image, they tend to improve subjective quality assessment which is at least correlated with comfort. They are also important because they introduce powerful depth mapping techniques, many of which can also be useful for improving viewing comfort as was shown earlier in this section.

Seam carving. This is a classic content retargeting method [155], which finds connected pixel paths in an image reaching from top to bottom or left to right (seams). Less important seams can then be removed in order to resize the image. Although it can lead to artefacts, seam carving has been influential, and several researchers have modified the classic seam carving retargeting algorithm to stereoscopic 3D. Jhou et al. extended this approach to RGBD images, so depth can also be resized to fit the target device [63]. Basha et al. introduced seam carving for pairs of stereo images. They exploited visibility relations between the two images (which pixels occlude which pixels in the corresponding view) in order to produce geometrically correct retargeting [12, 13]. They proved that the resulting pair is geometrically consistent with a feasible 3D scene.

Image Warping. A popular alternative to seam carving is to use mesh-based warping to transform an image. In order to extend this idea to stereo images, Chang et al. [21] proposed a method based on sparsely matched SIFT features followed by mesh warping. They preserve object aspect ratio and apply a linear depth mapping to achieve a range suitable for the target device, where a user can interactively choose the required depth. Yan et al. extend this idea by enforcing spatial

and temporal coherence of nearby features and the consistency of lines and planes in videos [201]. Zellinger et al. showed that these results can be further improved by linear optimisation [206].

These methods work well in regions where there are many interest points, but may introduce depth distortions in large featureless regions. Yoo et al. addressed this by introducing a novel energy function designed to ensure consistent depth for salient objects [203]. The warping is further refined by Liu et al. [111], who enforce spatial and temporal coherence of both disparity values and disparity variation. An alternative approach by Tasli and Alatan proposed user-assisted depth remapping based on superpixel primitives as graph nodes in a Markov Random Field [180]. Li et al. balance shape-preservation and depth-preservation constraints in their stereo retargeting model to derive warping functions which lead to natural-looking images, but they do not consider or measure viewing comfort [99, 100].

View interpolation. Another common way to produce new stereo pairs involves interpolation between existing views. An early example of this work is the 4D function called *the Lumigraph* which allows the construction of arbitrary new views of a static scene from a set of existing ones [45]. Zitnick et al. extended this principle to stereoscopic video [216]. Both methods build a 3D model of the scene, and Zitnick’s approach makes use of layered depth images to improve inpainting. Similarly, Bleyer et al. calculate a novel right view given a left view and a disparity map which can be helpful for producing disparity adjustments [16]. Smolic et al. represent the images using two boundary layers and one reliable layer [168] and use a combination of warping and additional hole-filling mechanisms to improve the resulting view. Although these methods made increasing use of depth information, they did not address depth scaling or discomfort. They are important because they lead to the multi-view method by Li et al. described in Sec. 4.1 [101].

A recent study by Chen et al. showed that 3D saliency models are useful for improving 3D retargeting. They used a 3D saliency model to improve both the seam carving and warping-based stereo retargeting methods [28].

5. Improved Displays

A stereoscopic display produces a simulation of a real 3D scene. True 3D displays are expected to solve most problems associated with stereo 3D [54], but everything seems to indicate that the near future of 3D displays will

be dominated by the stereoscopic technologies. This has driven recent research into improved stereoscopic displays, aimed at reducing perceived discomfort. Different types of displays suffer from different problems. For example, shutter glasses have been found to cause more discomfort than circularly polarised displays [211], but the latter are more likely to suffer from crosstalk and reduced contrast. Consequently, new display technology tends to address specialised problems inherent to each particular type of display.

5.1. *Adjusting for viewer pose*

An intermittent source of discomfort is the parallax distortion effect, which occurs when the viewer adjusts the position of their head. Jones et al. addressed this in their system [68] which was capable of calculating new views based on the actual viewing position determined by a face tracker. Different viewing positions also generate distortions due to affine transformations inconsistent with natural viewing. Li et al. did not address the issue of parallax, but instead proposed a model which can adjust the image on an anaglyph display for multiple fixed viewing positions. They demonstrate the principle using an anaglyph display and six viewing positions [104].

5.2. *Multi-focal displays*

Disparity mapping can reduce discomfort caused by the vergence-accommodation conflict, but only a novel display can address the underlying cause. Varifocal displays capable of automatic adjustment of the focal plane have been shown to improve comfort [166, 162], but they rely on a fast and accurate refocusing mechanism and accurate eye tracking and suffer from current technical limitations.

Multifocal displays such as the early prototype by Rolland et al. [146] do not rely on eye tracking. They stack multiple physical display planes to achieve near-correct focus cues. Akeley et al. were the first to explore displays with multiple focal distances [7] in order to reduce viewing discomfort. They reduced Rolland's 14 layers to only 3. An analysis by McKenzie et al. found that the maximum separation of the planes should be 1 Diopter [117]. A follow-up study found that accommodation responses to real and depth-filtered stimuli were equivalent for image-plane separations of 0.6 to 0.9 Diopter [116], and concluded that depth-filtering approaches based on multiple viewing planes can be used to precisely match accommodation and vergence.

Hoffman et al. showed that such a display is more comfortable than classic stereoscopic displays [52] and

that it produces natural-looking blur. Since they compared stereoscopic viewing to natural viewing, this work is considered the first one to conclusively prove that vergence-accommodation conflict is an actual cause of discomfort. Instead of multiple physical projection planes, the device by Love et al. [113] relied on a fast switchable lens with multiple focal states, making it a better candidate for miniaturisation.

Multi-focal displays require a static viewing position. This makes them unsuitable for e.g. cinema viewing, but interesting for head-mounted displays. Hoffman's device was very large and required the viewer's eyes to remain fixed, but Liu et al. introduced a see-through head-mounted display with multiple focal planes in the same year [110]. Their monocular prototype was aimed at augmented-reality applications, such that virtual objects would have the correct amount of focal blur, once inserted into a scene.

5.3. *Crosstalk*

Because crosstalk strongly depends on the display itself, much research has gone into hardware improvements, or combinations of hardware and software. Some progress is automatic: faster LCD switching times, improved materials, and better manufacturing have already reduced some of the causes of ghosting without targeting it specifically. In this section, we list some recent display advances specifically designed to reduce or eliminate crosstalk.

Several authors have explored projection-type autostereoscopic 3D display systems which utilise parallax barriers [142, 96, 105] in order to reduce crosstalk. Xue et al. reduced crosstalk in autostereoscopic displays by applying advanced backlight control [200]. Ma et al. combined image processing with a novel display which split a pixel into multiple zones [115]. Kim et al. [82] applied a micro lens array film and a mapping algorithm for a handheld diamond pentile-based display.

Many modern stereoscopic displays are typically incompatible with 2D viewing. A viewer without special equipment (passive or active glasses) will perceive a distorted combination of both the left and right view. Scher et al. proposed a method for reducing ghosting for 2D viewing by adding a third image, designed to cancel one of the original stereoscopic views. Viewers with 3D glasses will be presented with the left and right views only. For viewers without special glasses, the sum of the three images will amount to the left view only, without interference from the right view [151].

With the increased competition in the 3D display market, many innovations have been patented. Recent

patented improvements claimed to reduce crosstalk include a slantwise strip parallax barrier [20], spatially modulating illumination beams [93], improved slit gratings [59], improved polarisation [184], and a retro-reflective display [42]. There have also been devices with improved shutter timing [128].

6. Discussion

There has clearly been much progress in reducing visual discomfort, but there is still a long way to go. While some of the algorithms listed certainly improve comfort, others are a matter of taste. Indeed, it is not always clear if there is a solution that will be preferred by all viewers. However, subject to these caveats, what can we say about the general minimisation of visual discomfort?

First of all, it is important to perform the acquisition properly. This means parallel cameras with a baseline similar to the average human interocular distance. The cameras should be properly calibrated and aligned to avoid vertical parallax, and images should be cropped to avoid showing objects not visible to both eyes. Toe-in configuration is considered best avoided [9]. These simple measures can eliminate some of the worst causes of discomfort, and avoid the need for complicated post-processing.

It is also important to keep on-screen disparities in check, by concentrating most of the scene within a 3D volume that is relatively compact in depth. There are rules of thumb available for film makers, and following these can make any subsequent processing easier, and minimise strong distortions caused by subsequent image warping and nonlinear mapping. Cuts should be used sparingly, and should be arranged so that large disparity jumps are avoided. This can be difficult because cuts are typically introduced during the editing process, after the material has already been filmed. Where this is not possible, a cross-fade could be used instead of a cut, or a dynamic disparity mapping added during post-processing to reduce the discrepancy in depth. Where possible, fast motion in depth should be avoided, because vergence is impeded by incorrect accommodation cues.

6.1. Important Features

There is an inherent difficulty in comparing the performance of different models of discomfort prediction. The lack of standardised datasets means that different methods based on different features (such as disparity and motion statistics) will typically regress to different opinion scores based on different experiments with

different input data, making a head-to-head comparison impossible, and a large-scale comparison of all presented methods is beyond the scope of this review. Development of standardised benchmarks will help compare different methods on equal footing.

Still, most computational methods of discomfort are compared to other existing methods on small experiments. This provides an indication of which image features are useful, and largely confirms the insights from the study of visual discomfort. Existing evaluation suggests that a combination of multiple cues performs better than relying on a single discomfort cue [90, 32, 31]. Giving more weight to features in salient regions has also shown to be useful [31], though accurate gaze prediction remains a difficult problem. Finally, models which incorporate perceptual insights such as Weber's law [29], winner-takes-all strategy [97] or models of disparity perception in humans [35] and neural disparity models [135] have shown superior performance.

6.2. State of the Art

There are effective methods for dealing with crosstalk and many recent improvements. Since crosstalk is ultimately caused by imperfect display equipment, this issue is closely related to the development of display technology, as discussed in Section 5. There has been a steady progress in this field which is expected to continue, even though many of the innovations are covered by patents.

Excellent results have been obtained in terms of disparity mapping, as discussed in Section 4.1. Many methods are capable of automatically processing existing stereo images and videos, and they were shown to reduce discomfort. Powerful warping and scaling methods were adapted from image processing to reduce image and video distortion. This continues to be a very active field, so further improvements are likely.

Much less promising results have been obtained regarding artificial depth-of-field. Current eye-tracking technology does not seem to be good enough to allow artificial blurring capable of fooling the human eye. Even though some experiments reported improved immersion, no experiment conclusively showed an improvement in comfort, and many reported the effect to be distracting and unpopular [38, 19]. However, there is evidence that artificial DoF can alleviate the vergence-accommodation conflict [183] and recent experiments show that a simple DoF model can decrease discomfort for some viewers in a VR scenario [19], so more progress can be expected as eye tracking technology continues improving.

The effect of 3D saliency is inconclusive. Although many algorithms built on saliency models, the actual role of saliency in these models is hard to measure. While saliency maps may indicate which areas of the image are more likely to be viewed, and this knowledge can be used to boost average numbers, discomfort is often caused by peripheral effects such as excessive disparities near screen edges. Additionally, reducing distortions in salient regions often introduces more distortion in less salient regions, even though some viewers will also look at these.

Comparatively little has been done to address motion specifically. Several mapping methods exist for dealing with fast motion in depth, and they were shown to improve viewer comfort and produce pleasing motion. But it remains unclear how to deal with long scenes that contain many moving objects.

The success of recent methods that jointly address the multiple causes of discomfort indicates the importance of a broad approach, as discussed in Section 3.6. However, it will be necessary to balance different solutions, because they can conflict: e.g. disparity mapping can relax the vergence-accommodation conflict, but lead to less natural-looking motion and scene geometry. This will mean development of better cost functions, and a way to adapt them to individual users. Since causes of discomfort are so varied, it is to be expected that algorithms will address many different aspects of discomfort in the future.

6.3. Open Issues and Future Steps

Since this is a new field, direct comparison of methods is complicated by the lack of a standard benchmark shared by all groups. The development of the datasets discussed in Sec. 3.8 will hopefully solve this problem. A related issue is that all methods listed in this review attempt to model and reduce the “absolute” level of discomfort associated with stereoscopic viewing, but not all discomfort is due to stereoscopic effects. Many natural situations are distinctly uncomfortable, including extremely close objects (extreme convergence) and quick changes in depth, so it is not surprising that viewing a stereoscopic representation of such scenes is uncomfortable. Comparing stereoscopic material with a real-life baseline would be helpful in establishing the limits of what is possible, and lead to a better understanding of what makes stereoscopic material different but this is difficult and has largely been sidestepped in existing literature. Among the few studies which attempted to do this was the work of Hoffman et al. [52] which established that vergence-accommodation conflict was a major source of discomfort.

Viewing discomfort is very important because it can degrade viewing experience, but it is necessary to balance discomfort against other factors. For example, improving depth perception and vergence-accommodation conflict can result in otherwise degraded and unnatural image. Similarly, most guidelines encourage relatively “flat” scenes which might appear bland and unappealing. Recent work in Quality Assessment and Quality of Experience has begun to incorporate measures of discomfort, leading to an overall quality assessment of stereoscopic images and videos, as discussed in Sec. 3.7. The associated datasets will be important in measuring the overall effect of each new method for reducing discomfort.

It is more difficult to balance discomfort with creative freedom. It is apparent that comfortable stereoscopic material must be strongly constrained in terms of disparity range, and type and speed of movement. Many directors find these constraints too limiting, especially knowing that most 3D movies are also shown in 2D, which is a much less constrained medium. Development of new methods which do not affect scene geometry, such as dynamic range and selective blur [19] could help to relax some of these constraints.

There is a fundamental limitation to most presented algorithms in that they attempt to find a universal, generalised measure of discomfort, yet it is known that discomfort depends on many factors, and varies from individual to individual. Also, people will perceive a scene differently in terms of viewing location and specific sequence of fixations. A small number of algorithms presented in Sec. 4.1 can adapt to specific individuals, but this is still a small field. Furthermore, it has been shown that top-down processes can improve attention models [6], so computational models will need to incorporate high-level semantics from a scene understanding system. No existing models currently use such information, but powerful scene understanding systems are available and could be combined with existing models [182, 56]. A major difficulty is that so much of today’s media is not personalised, but viewed together with others, e.g. 3D movies in cinemas. This is also an important factor limiting the effectiveness of models based on eye tracking (in addition to latency and precision). The perceptual disparity model by Didyk et al. [35] can be used to process the image using special viewing hardware, modifying the disparities in a user-specific way, and this could potentially be extended to multiple viewers.

Personalised discomfort reduction is likely to accompany the spread of alternative viewing devices such as smartphones, tablets and HMDs. These devices are typically used by one user at a time, making personalised

methods more applicable. Additionally, since the viewing distance is not fixed in the case of hand-held devices, these algorithms will have to take changing viewing position into account.

7. Conclusions

We have presented a comprehensive overview of methods and techniques for reducing visual discomfort, in relation to stereoscopic 3D viewing. To our knowledge, this is the first review that specifically addresses solutions, rather than causes.

While many problems remain unsolved, there has been much progress in this interdisciplinary field. A combination of best practices during acquisition, together with recent post-processing algorithms, can significantly improve the 3D viewing experience.

Acknowledgement. This work was supported by the EPSRC grant EP/M01469X/1, “Geometric Evaluation of Stereoscopic Video”.

References

- [1] Ivy lab stereoscopic 3d image database for visual discomfort prediction, 2013. Online. Available: <http://ivylab.kaist.ac.kr/demo/3DVCA/3DVCA.htm>.
- [2] IEEE standard for quality of experience (QoE) and visual-comfort assessments of three-dimensional (3D) contents based on psychophysical studies. *IEEE Std 3333.1.1-2015*, pages 1–46, July 2015.
- [3] IEEE standards association stereoscopic (3d imaging) database, 2015. Online. Available: <http://grouper.ieee.org/groups/3dhf>.
- [4] 3d visual discomfort database, 2016. Online. Available: <https://www.scch.at/en/id-3d-visual-discomfort-database>.
- [5] P. Le Callet A. Benoit, P. Campisi, and R. Cousseau. Quality assessment of stereoscopic images. *EURASIP Journal on Image and Video Processing*, 659024, 2008.
- [6] D. Sihite A. Borji and L. Itti. Computational modeling of top-down visual attention in interactive environments. In *Proc. BMVC*, pages 85.1–85.12, May 2011.
- [7] K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks. A stereo display prototype with multiple focal distances. *ACM Trans. on Graphics*, 23:804813, 2004.
- [8] T. Bando, A. Iijima, and S. Yano. Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: A review. *Displays*, 33(2):76–83, 2012.
- [9] M. S. Banks, J. C. A. Read, R. S. Allison, and S. J. Watt. Stereoscopia and the human visual system. *SMPTE Motion Imaging Journal*, 121(4):24–43, May 2012.
- [10] M. Barkowsky, J. Li, T. Han, S. Youn, J. Ok, C. Lee, C. Hedberg, I. V. Ananth, K. Wang, K. Brunnström, and P. Le Callet. Towards standardized 3DTV QoE assessment: cross-lab study on display technology and viewing environment parameters. In *Proc. SPIE 8648, Stereoscopic Displays and Applications XXIV*, volume 8648, March 2013.
- [11] M. Barreda-Angeles, R. Pepion, E. Bosc, P. Le Callet, and A. Pereda-Banos. Exploring the effects of 3D visual discomfort on viewers’ emotions. In *IEEE International Conference on Image Processing (ICIP)*, pages 753–757, October 2014.
- [12] T. Basha, Y. Moses, and S. Avidan. Geometrically consistent stereo seam carving. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1816–1823, November 2011.
- [13] T. D. Basha, Y. Moses, and S. Avidan. Stereo seam carving a geometrically consistent approach. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 35(10):2513–2525, October 2013.
- [14] M. Bernhard, C. Dell’mour, M. Hecher, E. Stavrakis, and M. Wimmer. The effects of fast disparity adjustment in gaze-controlled stereoscopic applications. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA ’14, pages 111–118, New York, 2014. ACM.
- [15] Y. Bi and J. Zhou. A visual comfort metric for stereoscopic 3D video based on SMDE approach. In *IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC)*, pages 72–77, August 2012.
- [16] M. Bleyer, M. Gelautz, C. Rother, and C. Rhemann. A stereo approach that handles the matting problem via image warping. In *IEEE International Conference on Pattern Recognition (CVPR)*, pages 501–508, 2009.
- [17] W. Blohm, I. P. Beldie, K. Schenke, K. Fazel, and S. Pastoor. Stereoscopic image representation with synthetic depth of field. *Journal of the Society for Information Display*, 5(3):307–313, 1997.
- [18] T. Blum, M. Wieczorek, A. Aichert, R. Tibrewal, and N. Navab. The effect of out-of-focus blur on visual discomfort when using stereo displays. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 13–17, October 2010.
- [19] K. Carnegie and T. Rhee. Reducing visual discomfort with HMDs using dynamic depth of field. *IEEE Computer Graphics and Applications*, 35(5):34–41, September 2015.
- [20] G. E. Casner, J. C. Schultz, M. J. Sykora, and B. E. Brooks. Methods of reducing perceived image crosstalk in a multiview display, December 2012. US Patent 8,339,333.
- [21] C. H. Chang, C. K. Liang, and Y. Y. Chuang. Content-aware display adaptation and interactive editing for stereoscopic images. *IEEE Trans. on Multimedia*, 13(4):589–601, August 2011.
- [22] Y. C. Chang, C. Y. Ma, and Y. P. Huang. Crosstalk suppression by image processing in 3D display. *SID Symposium Digest of Technical Papers*, 41(1):124–127, 2010.
- [23] A. Chappuis, M. Rerabek, P. Hanhart, and T. Ebrahimi. Subjective evaluation of an active crosstalk reduction system for mobile autostereoscopic displays. In *Proc. SPIE 9011, Stereoscopic Displays and Applications XXV*, volume 9011, March 2014.
- [24] L. G. Chen, C. Wu, C. T. Li, Y. C. Lai, C. C. Cheng, and L. H. Huang. Nonlinear depth remapping system and method thereof, November 2012. US Patent App. 13/112,854.
- [25] M. J. Chen, D. K. Kwon, L. K. Cormack, and A. C. Bovik. Optimizing 3D image display using the stereoacuity function. In *IEEE International Conference on Image Processing (ICIP)*, pages 617–620, Sept 2012.
- [26] M. J. Chen, C. C. Su, D. K. Kwon, L. K. Cormack, and A. C. Bovik. Full-reference quality assessment of stereopairs accounting for rivalry. *Signal Processing: Image Communication*, 28(9):1143–1155, October 2013.
- [27] W. Chen, J. Fournier, M. Barkowsky, and P. Le Callet. New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone. In *Proc. SPIE*

- 7863, *Stereoscopic Displays and Applications XXII*, volume 7863, February 2011.
- [28] Y. Chen, Y. Pan, M. Song, and M. Wang. Image retargeting with a 3D saliency model. *Signal Processing*, 112:53–63, July 2015.
 - [29] Z. Chen, J. Shi, X. Huang, L. Yun, and Y. Tai. Visual comfort modeling for disparity in 3D contents based on Weber-Fechner’s law. *Journal of Display Technology*, 10(12):1001–1009, December 2014.
 - [30] S. H. Cho and H. B. Kang. A new visual comfort-based stereoscopic image retargeting method. In *Advances in Visual Computing*, volume 8033 of *Lecture Notes in Computer Science*, pages 290–300. Springer Berlin Heidelberg, 2013.
 - [31] S. H. Cho and H. B. Kang. Prediction of visual discomfort in watching 3D video using multiple features. In *IEEE South-west Symposium on Image Analysis and Interpretation (SSIAI)*, pages 65–68, April 2014.
 - [32] J. Choi, D. Kim, S. Choi, and K. Sohn. Visual fatigue modeling and analysis for stereoscopic video. *Optical Engineering*, 51(1):017206–1–017206–11, 2012.
 - [33] D. Corrigan, F. Pitié, V. Morris, A. Rankin, M. Linnane, G. Kearney, M. Gorzel, M. O’Dea, C. Lee, and A. Kokaram. A video database for the development of stereo-3d post-production algorithms. In *Conference on Visual Media Production (CVMP)*, pages 64–73, November 2010.
 - [34] S. Delis, N. Nikolaidis, and I. Pitas. Automatic detection of depth jump cuts and bent window effects in stereoscopic videos. In *IEEE IVMSWP Workshop*, pages 1–4, June 2013.
 - [35] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, and H. P. Seidel. A perceptual model for disparity. *ACM Trans. on Graphics (Proceedings SIGGRAPH 2011, Vancouver)*, 30(4), 2011.
 - [36] C. Doutre and P. Nasiopoulos. Crosstalk cancellation in 3D video with local contrast reduction. In *European Signal Processing Conference*, pages 1884–1888, August 2011.
 - [37] C. Doutre and P. Nasiopoulos. Optimized contrast reduction for crosstalk cancellation in 3D displays. In *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, pages 1–4, May 2011.
 - [38] A. T. Duchowski, D. H. House, J. Gestring, R. I. Wang, K. Krejtz, I. Krejtz, R. Mantiuk, and B. Bazyluk. Reducing visual discomfort of 3D stereoscopic displays with gaze-contingent depth-of-field. In *Proceedings of the ACM Symposium on Applied Perception*, SAP ’14, pages 39–46, New York, NY, USA, 2014. ACM.
 - [39] I. Feldmann, O. Schreer, and P. Kauff. Nonlinear depth scaling for immersive video applications. In *WIAMIS*, 2003.
 - [40] The Foundry. Ocula, 2010. <http://www.thefoundry.co.uk/>.
 - [41] G. Fry. Further experiments on the accommodative convergence relationship. *American Journal of Optometry*, 16:325–334, 1939.
 - [42] C. Gao and J. Xiao. Methods to eliminate/reduce the crosstalk artifacts of the retro-reflective auto-stereoscopic 3D display, April 2013. US Patent 8,411,135.
 - [43] L. Goldmann, F. De Simone, and T. Ebrahimi. A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video. In *Electronic Imaging, 3D Image Processing and Applications*, 2010.
 - [44] L. Goldmann, F. De Simone, and T. Ebrahimi. Impact of acquisition distortion on the quality of stereoscopic images. In *International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, Scottsdale, January 2010.
 - [45] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *SIGGRAPH*, pages 43–54, 1996.
 - [46] K. Ha and M. Kim. A perceptual quality assessment metric using temporal complexity and disparity information for stereoscopic video. In *IEEE International Conference on Image Processing (ICIP)*, pages 2525–2528, September 2011.
 - [47] J. M. Harris and S. N. Watamaniuk. Speed discrimination of motion-in-depth using binocular cues. *Vision Research*, 35(7):885–896, 1995.
 - [48] S. He, T. Zhang, and D. Doyen. Visual discomfort prediction for stereo contents. *Proc. SPIE*, 7863:78631X–78631X–9, 2011.
 - [49] S. Heinzle, P. Greisen, D. Gallup, C. Chen, D. Saner, A. Smolic, A. Burg, W. Matusik, and M. Gross. Computational stereo camera system with programmable control loop. *ACM Transactions on Graphics*, 30:1–10, 2011.
 - [50] S. Hillaire, A. Lecuyer, R. Cozot, and G. Casiez. Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. In *IEEE Virtual Reality Conference*, pages 47–50, Reno, March 2008.
 - [51] M. Hirsch, D. Lanman, G. Wetzstein, and R. Raskar. Tensor displays. In *ACM SIGGRAPH 2012 Emerging Technologies*, pages 24:1–24:1, New York, NY, USA, 2012. ACM.
 - [52] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3:33):130, 2008.
 - [53] D. M. Hoffman, V. I. Karasev, and M. S. Banks. Temporal presentation protocols in stereoscopic displays: Flicker visibility, perceived motion, and perceived depth. *Journal of the SID*, 19(3):255–281, 2011.
 - [54] N. S. Holliman, N. A. Dodgson, G. E. Favalora, and L. Pockett. Three-dimensional displays: A review and applications analysis. *IEEE Trans. on Broadcasting*, 57(2):362–371, June 2011.
 - [55] H. K. Hong. Reduction of spatially non-uniform 3D crosstalk for stereoscopic display using shutter glasses. *Displays*, 33(3):136–141, July 2012.
 - [56] L. Hotz, B. Neumann, K. Terzić, and J. Šochman. Feedback between low-level and high-level image processing. Technical Report Report FBI-HH-B-278/07, Universität Hamburg, Hamburg, 2007.
 - [57] P. A. Howarth. Potential hazards of viewing 3-d stereoscopic television, cinema and computer games: a review. *Ophthalmic & Physiological Optics*, 31:111–122, 2011.
 - [58] K. Huang, K. Lee, and H. Lin. Crosstalk issue in stereo/autostereoscopic display. In *International Display Manufacturing Conference*, pages 2–18, 2009.
 - [59] X. Huang and F. Li. Slit grating and display device, January 2014. US Patent App. 13/941,681.
 - [60] A. D. Hwang and E. Peli. Instability of the perceived world while watching 3D stereoscopic imagery: A likely source of motion sickness symptoms. *i-Perception*, 5(6):515–535, October 2014.
 - [61] I. Iatsun, M. C. Larabi, and C. Fernandez-Maloigne. Investigation and modeling of visual fatigue caused by S3D content using eye-tracking. *Displays*, 39:11–25, October 2015.
 - [62] I. Ideses and L. Yaroslavsky. Three methods that improve the visual quality of colour anaglyphs. *Journal of Optics A: Pure and Applied Optics*, 7:755–762, 2005.
 - [63] W. C. Jhou, Y. H. Lin, and J. L. Wu. Seam carving for color-plus-depth 3D image. In *IEEE International Symposium on Multimedia (ISM)*, pages 82–85, December 2014.
 - [64] Q. Jiang and F. Shao. 3D visual comfort assessment via sparse coding. In Y. J. Zhang, editor, *Image and Graphics*, volume 9217 of *Lecture Notes in Computer Science*, pages 10–20. Springer International Publishing, 2015.
 - [65] Q. Jiang, F. Shao, G. Jiang, M. Yu, and Z. Peng. Three-dimensional visual comfort assessment via preference learn-

- ing. *Journal of Electronic Imaging*, 24(4):043002, July 2015.
- [66] Q. Jiang, F. Shao, G. Jiang, M. Yu, Z. Peng, and C. Yu. A depth perception and visual comfort guided computational model for stereoscopic 3D visual saliency. *Signal Processing: Image Communication*, 38:57–69, October 2015.
- [67] E. W. Jin, M. E. Miller, S. Endrikhovski, and C. D. Cerosaletti. Stereoscopic display system with flexible rendering of disparity map according to the stereoscopic fusing capability of the observer, January 2012. US Patent 8,094,927.
- [68] G. R. Jones, D. Lee, N. S. Holliman, and D. Ezra. Controlling perceived depth in stereoscopic images. In *Stereoscopic Displays and Virtual Reality Systems VIII, Proceedings of the SPIE*, volume 4297, pages 42–53, June 2001.
- [69] S. W. Jung and S. J. Ko. Depth sensation enhancement using the just noticeable depth difference. *IEEE Trans. on Image Processing*, 21(8):3624–3637, August 2012.
- [70] Y. Jung, H. Sohn, S. I. Lee, and Y. Ro. Visual comfort improvement in stereoscopic 3D displays using perceptually plausible assessment metric of visual comfort. *IEEE Trans. on Consumer Electronics*, 60(1):1–9, February 2014.
- [71] Y. J. Jung, H. Sohn, S. I. Lee, H. W. Park, and Y. M. Ro. Predicting visual discomfort of stereoscopic images using human attention model. *IEEE Trans. on Circuits and Systems for Video Technology*, 23(12):2077–2082, Dec 2013.
- [72] Y. J. Jung, H. Sohn, S. I. Lee, F. Speranza, and Y. M. Ro. Visual importance- and discomfort region-selective low-pass filtering for reducing visual discomfort in stereoscopic displays. *IEEE Trans. on Circuits and Systems for Video Technology*, 23(8):1408–1421, Aug 2013.
- [73] Y. J. Jung, H. Sohn, and Y. M. Ro. Visual discomfort visualizer using stereo vision and time-of-flight depth cameras. *IEEE Trans. on Consumer Electronics*, 58(2):246–254, May 2012.
- [74] D. H. Kang, E. J. Lee, J. H. Lee, and J. K. Song. Perceptual strength of 3-D crosstalk in both achromatic and color images in stereoscopic 3-D displays. *IEEE Trans. on Image Processing*, 21(7):3253–3261, July 2012.
- [75] P. Kellnhofer, T. Ritschel, K. Myszkowski, and H. P. Seidel. Optimizing disparity for motion in depth. In *Eurographics Symposium on Rendering 2013*, volume 32, 2013.
- [76] P. Kellnhofer, T. Ritschel, K. Myszkowski, and H. P. Seidel. Improving perception of binocular stereo motion on 3D display devices. In *Proc. SPIE 9011, Stereoscopic Displays and Applications XXV*, volume 9011, March 2014.
- [77] L. J. Kerofsky, S. G. Deshpande, and S. J. Daly. System for crosstalk reduction, March 2012. US Patent App. 12/878,644.
- [78] D. Kim, S. Choi, S. Park, and K. Sohn. Stereoscopic visual fatigue measurement based on fusional response curve and eyeblinks. In *International Conference on Digital Signal Processing (DSP)*, pages 1–6, July 2011.
- [79] D. Kim, S. Choi, and K. Sohn. Visual comfort enhancement for stereoscopic video based on binocular fusion characteristics. *IEEE Trans. on Circuits and Systems for Video Technology*, 23(3):482–487, March 2013.
- [80] D. Kim and K. Sohn. Visual fatigue prediction for stereoscopic image. *IEEE Trans. on Circuits and Systems for Video Technology*, 21(2):231–236, February 2011.
- [81] H. Kim, S. Lee, and A. C. Bovik. Saliency prediction on stereoscopic videos. *IEEE Transactions on Image Processing*, 23(4):1476–1490, April 2014.
- [82] J. Kim, C. K. Lee, Y. Jeong, C. Jang, J. Y. Hong, W. Lee, Y.C. Shin, J. H. Yoon, and B. Lee. Crosstalk-reduced dual-mode mobile 3D display. *Journal of Display Technology*, 11(1):97–103, January 2015.
- [83] M. Kim, S. Lee, C. Choi, G. M. Um, N. Hur, and J. Kim. Depth scaling of multiview images for automultiscopic 3D monitors. In *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video*, pages 181–184, May 2008.
- [84] T. Kim, S. Lee, and A. C. Bovik. Transfer function model of physiological mechanisms underlying temporal visual discomfort experienced when viewing stereoscopic 3D images. *IEEE Trans. on Image Processing*, 24(11):4335–4347, November 2015.
- [85] S. Kitrosser. Photography in the service of stereoscopy. *Journal of Imaging Science and Technology*, 42:295–300, July 1998.
- [86] D. Kondermann, R. Nair, S. Meister, W. Mischler, B. Gössefeld, K. Honauer, S. Hofmann, C. Brenner, and B. Jähne. Stereo ground truth with error bars. In D. Cremers, I. Reid, H. Saito, and M. H. Yang, editors, *Computer Vision – ACCV 2014*, volume 9007 of *Lecture Notes in Computer Science*, pages 595–610. Springer International Publishing, 2015.
- [87] J. Konrad, B. Lacotte, and E. Dubois. Cancellation of image crosstalk in time-sequential displays of stereoscopic video. *IEEE Trans. on Image Processing*, 9(5):897–908, May 2000.
- [88] F. L. Kooi and A. Toet. Visual comfort of binocular and 3-D displays. *Displays*, 25:99108, 2004.
- [89] S. J. Koppal, C. L. Zitnick, M. K. Cohen, S. B. Kang, B. Ressler, and A. Colburn. A viewer-centric editor for 3d movies. *IEEE Computer Graphics and Applications*, 31(1):20, 2011.
- [90] M. Lambooi, W. Ijsselstein, M. Fortuin, and I. Heynderickx. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, 53(3):114, 2009.
- [91] M. Lambooi, W. A. Ijsselstein, and I. Heynderickx. Visual discomfort of 3D TV: Assessment methods and modeling. *Displays*, 32(4):209 – 218, 2011. Visual Image Safety.
- [92] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross. Nonlinear disparity mapping for stereoscopic 3D. *ACM Trans. on Graphics*, 29(3):10, 2010.
- [93] K. Lee, K. C. Huang, and C. H. Tsai. Stereo-image displaying apparatus and method for reducing stereo-image cross-talk, March 2013. US Patent 8,403,487.
- [94] K. Lee and S. Lee. 3d perception based quality pooling: Stereopsis, binocular rivalry, and binocular suppression. *IEEE J. of Selected Topics in Signal Processing*, 9(3):533–545, April 2015.
- [95] K. Lee, A. K. Moorthy, S. Lee, and A. C. Bovik. 3d visual activity assessment based on natural scene statistics. *IEEE Transactions on Image Processing*, 23(1):450–465, January 2014.
- [96] K. H. Lee, Y. Park, H. Lee, S. K. Yoon, and S. K. Kim. Crosstalk reduction in auto-stereoscopic projection 3D display system. *Optics Express*, 20(18):19757–19768, Aug 2012.
- [97] S. I. Lee, Y. J. Jung, H. Sohn, and Y. M. Ro. Experimental investigation of discomfort combination: toward visual discomfort prediction for stereoscopic videos. *Journal of Electronic Imaging*, 23(1), January 2014.
- [98] L. Leroy, P. Fuchs, and G. Moreau. Visual fatigue reduction for immersive stereoscopic displays by disparity, content, and focus-point adapted blur. *IEEE Trans. on Industrial Electronics*, 59(10):3998–4004, October 2012.
- [99] B. Li, L. Duan, C. W. Lin, and W. Gao. Region-based depth-preserving stereoscopic image retargeting. In *IEEE International Conference on Image Processing (ICIP)*, pages 2903–2907, October 2014.
- [100] B. Li, L. Y. Duan, C. W. Lin, T. Huang, and W. Gao. Depth-preserving warping for stereo image retargeting. *IEEE Trans. on Image Processing*, 24(9):2811–2826, September 2015.
- [101] D. Li, X. Qiao, D. Zang, L. Wang, and M. Zhang. On adjustment of stereo parameters in multiview synthesis for planar 3D

- displays. *Journal of the Society for Information Display*, 2015.
- [102] D. Li, D. Zang, X. Qiao, L. Wang, and M. Zhang. 3D synthesis and crosstalk reduction for lenticular autostereoscopic displays. *Journal of Display Technology*, 11(11):939–946, November 2015.
- [103] J. Li, M. Barkowsky, and P. Le Callet. Visual discomfort in 3DTV: Definitions, causes, measurement, and modeling. In A. Kondoz and T. Dagiuklas, editors, *Novel 3D Media Technologies*, pages 185–209. Springer New York, 2015.
- [104] Q. Li and D. Schonfeld. General stereoscopic distortion rectification due to arbitrary viewer motion in binocular stereoscopic display. In *Proc. SPIE 9011, Stereoscopic Displays and Applications XXV*, volume 9011, March 2014.
- [105] J. L. Liang, Q. Wang, S. F. Zang, W. X. Zhao, and Q. H. Wang. Autostereoscopic 3D projection display with low crosstalk. *SID Symposium Digest of Technical Papers*, 46(1):1424–1426, 2015.
- [106] H. S. Lin, S. H. Guan, C. T. Lee, and M. Ouhyoung. Stereoscopic 3D experience optimization using cropping and warping. In *SIGGRAPH Asia 2011 Sketches*, SA '11, pages 40:1–40:2, New York, 2011. ACM.
- [107] J. S. Lipscomb and W. L. Wooten. Reducing crosstalk between stereoscopic views. In *Proc. SPIE 2177, Stereoscopic Displays and Virtual Reality Systems*, April 1994.
- [108] L. Lipton. *Stereographics, developers handbook*, 1997.
- [109] C. W. Liu, T. H. Huang, M. H. Chang, K. Y. Lee, C. K. Liang, and Y. Y. Chuang. 3D cinematography principles and their applications to stereoscopic media processing. In *ACM international conference on Multimedia*, pages 253–262, Scottsdale, November 2011.
- [110] S. Liu, D. Cheng, and H. Hua. An optical see-through head mounted display with addressable focal planes. In *International Symposium on Mixed and Augmented Reality*, page 3341. IEEE, 2008.
- [111] Y. Liu, L. Sun, and S. Yang. A retargeting method for stereoscopic 3D video. *Computational Visual Media*, 1(2):119–127, 2015.
- [112] J. López, J. Rodrigo, D. Jiménez, and J. Méndez. Stereoscopic 3D video quality assessment based on depth maps and video motion. *EURASIP Journal on Image and Video Processing*, 2013(1), 2013.
- [113] G. D. Love, D. M. Hoffman, P. J. W. Hands, J. Gao, A. K. Kirby, and M. S. Banks. High-speed switchable lens enables the development of a volumetric stereoscopic display. *Optics Express*, 17:1571515725, 2009.
- [114] Schor C. M. and J. Kotulak. Dynamic interactions between accommodation and convergence are velocity sensitive. *Vision Research*, 26:927–942, 1986.
- [115] C. Y. Ma, Y. C. Chang, and Y. P. Huang. Multi-zone digital crosstalk reduction by image processing in 3D display. *Journal of Display Technology*, 10(6):488–493, June 2014.
- [116] K. J. MacKenzie, R. A. Dickson, and S. J. Watt. Vergence and accommodation to multiple-image-plane stereoscopic displays: “real world” responses with practical image-plane separations? *J. Electron. Imaging*, 21(1), February 2012.
- [117] K. J. MacKenzie, D. M. Hoffman, and S. J. Watt. Accommodation to multiple-focal-planes displays: implications for improving stereoscopic displays, and for accommodation control. *Journal of Vision*, 10(8):22, 2010.
- [118] S. Mangiat and J. Gibson. Disparity remapping for handheld 3D video communications. In *IEEE International Conference on Emerging Signal Processing Applications (ESPA)*, pages 147–150, January 2012.
- [119] B. Masia, G. Wetzstein, C. Aliaga, R. Raskar, and D. Gutierrez. Display adaptive 3D content remapping. *Computers & Graphics*, 37(8):983–996, December 2013.
- [120] D. F. McAllister, Y. Zhou, and S. Sullivan. Methods for computing color anaglyphs. pages 75240S–75240S–12, 2010.
- [121] J. P. McIntire, P. R. Havig, and E. E. Geiselman. Stereoscopic 3D displays and human performance: A comprehensive review. *Displays*, 35(1):18–26, January 2014.
- [122] L. Meesters, W. IJsselstein, and P. Seuntjens. A survey of perceptual evaluations and requirements of three-dimensional TV. *IEEE Trans. Circuits and Systems for Video Technology*, 14(3):381–391, 2004.
- [123] B. Mendiburu. *3D Movie Making: Stereoscopic digital cinema from script to screen*. Focal Press, 2009.
- [124] A. Mittal, A. K. Moorthy, J. Ghosh, and A. C. Bovik. Algorithmic assessment of 3D quality of experience for images and videos. In *Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE)*, pages 338–343, January 2011.
- [125] A. K. Moorthy and A. C. Bovik. A survey on 3D quality of experience and 3D quality assessment. In *Proc. SPIE 8651, Human Vision and Electronic Imaging XVIII*, volume 8651, March 2013.
- [126] A. K. Moorthy, C. C. Su, A. Mittal, and A. C. Bovik. Subjective evaluation of stereoscopic image quality. *Signal Processing: Image Communication*, pages 870–883, 2013.
- [127] T. J. Mu, J. J. Sun, R. R. Martin, and S. M. Hu. A response time model for abrupt changes in binocular disparity. *The Visual Computer*, 31(5):675–687, 2015.
- [128] Y. Nakahata and M. Nakagawa. Image display apparatus, image display observation system, and image display method, October 2013. US Patent 8,564,646.
- [129] Y. Nazzar, J. Bouchard, and J. J. Clark. Detection of stereo window violation in 3D movies. In *ACM SIGGRAPH 2014 Posters*, SIGGRAPH '14, pages 61:1–61:1, New York, 2014. ACM.
- [130] Y. Nojiri, H. Yamanoue, A. Hanazato, and F. Okana. Measurement of parallax distribution, and its application to the analysis of visual comfort for stereoscopic HDTV. *Proc. SPIE*, 5006:195–205, 2003.
- [131] C. Oh, B. Ham, S. Choi, and K. Sohn. Visual fatigue relaxation for stereoscopic video via nonlinear disparity remapping. *IEEE Trans. on Broadcasting*, 61(2):142–153, June 2015.
- [132] H. Oh, S. Lee, and A. C. Bovik. Stereoscopic 3d visual discomfort prediction: A dynamic accommodation and vergence interaction model. *IEEE Trans. on Image Processing*, 25(2):615–629, February 2016.
- [133] D. Pajak, R. Herzog, R. Mantiuk, P. Didyk, E. Eisemann, K. Myszkowski, and K. Pulli. Perceptual depth compression for stereo applications. *Computer Graphics Forum*, 33(2):195–204, 2014.
- [134] J. Park, S. Lee, and A. C. Bovik. 3D visual discomfort prediction: Vergence, foveation, and the physiological optics of accommodation. *IEEE Journal of Selected Topics in Signal Processing*, 8(3):415–427, June 2014.
- [135] J. Park, H. Oh, S. Lee, and A. C. Bovik. 3D visual discomfort predictor: Analysis of disparity and neural activity statistics. *IEEE Trans. on Image Processing*, 24(3):1101–1114, March 2015.
- [136] J. Y. Park, Y. T. Kim, G. Y. Sung, and D. K. Nam. Apparatus and method for inter-view crosstalk reduction, April 2011. US Patent App. 12/662,893.
- [137] M. Park, J. Luo, and A. C. Gallagher. Toward assessing and improving the quality of stereo images. *IEEE Journal of Selected Topics in Signal Processing*, 6(5):460–470, September 2012.
- [138] M. C. Park and S. Mun. Overview of measurement methods for

- factors affecting the human visual system in 3D displays. *Journal of Display Technology*, 11(11):877–888, November 2015.
- [139] A. S. Percival. The relation of convergence. *Ophthalmological Review*, 11:313–328, 1892.
- [140] A. F. Perrin, H. Xu, E. Kroupi, M. Řeřábek, and T. Ebrahimi. Multimodal dataset for assessment of quality of experience in immersive multimedia. In *ACM Multimedia Conference*, 2015.
- [141] Y. Pritch, M. Ben-Ezra, and S. Peleg. Automatic disparity control in stereo panoramas (omnistereo). *OMNIVIS*, 2000.
- [142] L. Qi, Q. H. Wang, J. Y. Luo, W. X. Zhao, and C. Q. Song. An autostereoscopic 3d projection display based on a lenticular sheet and a parallax barrier. *Journal of Display Technology*, 8(7):397–400, July 2012.
- [143] S. Qi and J. Ho. Shift-map based stereo image retargeting with disparity adjustment. In K. M. Lee, Y. Matsushita, J. M. Rehg, and Z. Hu, editors, *Computer Vision ACCV 2012*, volume 7727 of *Lecture Notes in Computer Science*, pages 457–469. Springer Berlin Heidelberg, 2013.
- [144] J. C. A. Read and I. Bohr. User experience while viewing stereoscopic 3D television. *Ergonomics*, 57(8):1140–1153, 2014.
- [145] C. Richardt, L. Świrski, I. P. Davies, and N. A. Dodgson. Predicting stereoscopic viewing comfort using a coherence-based computational model. In *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging*, CAe '11, pages 97–104, New York, NY, USA, 2011. ACM.
- [146] J. P. Rolland, M. W. Krueger, and A. Goon. Multifocal planes head-mounted displays. *Applied Optics*, 39:3209–3215, July 2000.
- [147] S. Ryu, D. H. Kim, and K. Sohn. Stereoscopic image quality metric based on binocular perception model. In *IEEE International Conference on Image Processing (ICIP)*, pages 609–612, September 2012.
- [148] T. Sakamoto and T. Yakoh. Design and evaluation of binocular stereoscopic system with variable vergence mechanism for near-field operation. *IEEJ Trans. on Industry Applications*, 134(9):792–800, 2014.
- [149] W. Sanders and D. F. McAllister. Producing anaglyphs from synthetic images. In *Proceedings of SPIE 5006*, pages 348–358, 2003.
- [150] H. Sanftmann and D. Weisskopf. Anaglyph stereo without ghosting. *Computer Graphics Forum*, 30(4):1251–1259, June 2011.
- [151] S. Scher, J. Liu, R. Vaish, P. Gunawardane, and J. Davis. 3D+2DTV: 3D displays with no ghosting for viewers without glasses. *ACM Trans. on Graphics*, 32(3):21:1–21:10, July 2013.
- [152] C. Schor. A dynamic model of cross-coupling between accommodation and convergencesimulations of step and frequency responses. *Optometry and Vision Science*, 69:258–269, 1992.
- [153] C. M. Schor. The relationship between fusional vergence eye movements and fixation disparity. *Vision Research*, 19(12):1359–1367, 1979.
- [154] C. M. Schor. The influence of interactions between accommodation and convergence on the lag of accommodation. *Ophthalmic and Physiological Optics*, 19(2):134–150, 1999.
- [155] V. Setlur, S. Takagi, R. Raskar, M. Gleicher, and B. Gooch. Automatic image retargeting. In *International Conference on Mobile and Ubiquitous Multimedia*, pages 59–68, 2005.
- [156] P. J.H. Seuntiëns, L. M. J. Meesters, and W. A. IJsselstein. Perceptual attributes of crosstalk in 3D images. *Displays*, 26(4-5):177–183, October 2005.
- [157] F. Shao, Z. Li, Q. Jiang, G. Jiang, M. Yu, and Z. Peng. Visual discomfort relaxation for stereoscopic 3D images by adjusting zero-disparity plane for projection. *Displays*, 39:125–132, October 2015.
- [158] F. Shao, W. Lin, S. Gu, G. Jiang, and T. Srikanthan. Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics. *IEEE Trans. on Image Processing*, 22(5):1940–1953, May 2013.
- [159] C. Sheard. The prescription of prisms. *American Journal of Optometry*, 11:364–378, 1934.
- [160] J. E. Sheedy, J. Hayes, and J. Engle. Is all asthenopia the same? *Optom. Vision Science*, 80:732–739, 2003.
- [161] S. Shestak, D. Kim, and Y. Kim. How much crosstalk can be allowed in a stereoscopic system at various grey levels? In *Proc. SPIE 8288, Stereoscopic Displays and Applications XXIII*, volume 8288, February 2012.
- [162] T. Shibata, T. Kawai, K. Ohta, M. Otsuki, N. Miyake, Y. Yoshihara, and T. Iwasaki. Stereoscopic 3-d display with optical correction for the reduction of the discrepancy between accommodation and convergence. *Journal of the Society for Information Display*, 13:665–671, August 2005.
- [163] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision*, 11(8:11):129, 2011.
- [164] T. Shibata, F. Muneyuki, K. Oshima, J. Yoshitake, and T. Kawai. Comfortable stereo viewing on mobile devices. In *Proc. SPIE 8648, Stereoscopic Displays and Applications XXIV*, volume 8648, March 2013.
- [165] N. Shibuhisa, K. Tokui, K. I. Iwauchi, and S. Saito. Study on reducing the cardboard effect for natural perception using adaptive disparity mapping. *SID Symposium Digest of Technical Papers*, 43(1):961–964, 2012.
- [166] S. Shiwa, K. Omura, and F. Kishino. Proposal for a 3-D display with accommodative compensation: 3DDAC. *Journal of the Society for Information Display*, 4:255–261, December 1996.
- [167] F. A. Smit, R. van Liere, and B. Froehlich. Non-uniform crosstalk reduction for dynamic scenes. In *IEEE Virtual Reality Conference*, pages 139–146, March 2007.
- [168] A. Smolic, K. Miller, K. Dix, P. Merkle, P. Kauff, and T. Wiegand. Intermediate view interpolation based on multiview video plus depth for advanced 3D video systems. In *IEEE International Conference on Image Processing (ICIP)*, pages 2448–2451, 2008.
- [169] H. Sohn, Y. J. Jung, S. I. Lee, H. W. Park, and Y. M. Ro. Investigation of object thickness for visual discomfort prediction in stereoscopic images. *Proc. SPIE*, 8288:82880Q–82880Q–8, 2012.
- [170] H. Sohn, Y. J. Jung, S. I. Lee, and Y. M. Ro. Predicting visual discomfort using object size and disparity information in stereoscopic images. *IEEE Trans. on Broadcasting*, 59(1):28–37, March 2013.
- [171] H. Sohn, Y. J. Jung, S. I. Lee, F. Speranza, and Y. M. Ro. Disparity remapping to ameliorate visual comfort of stereoscopic video. 8648, March 2013.
- [172] H. Sohn, Y. J. Jung, S. I. Lee, F. Speranza, and Y. M. Ro. Visual comfort amelioration technique for stereoscopic images: Disparity remapping to mitigate global and local discomfort causes. *IEEE Trans. on Circuits and Systems for Video Technology*, 24(5):745–758, May 2014.
- [173] H. Sohn, Y. J. Jung, and Y. M. Ro. Crosstalk reduction in stereoscopic 3D displays: Disparity adjustment using crosstalk visibility index for crosstalk cancellation. *Optics Express*, 22(3):3375–3392, February 2014.
- [174] H. Sohn, Y. J. Jung, and Y. M. Ro. Local disparity remapping to enhance depth quality of stereoscopic 3D images using stereoacuity function. In *Proc. SPIE 9011, Stereoscopic Displays and Applications XXV*, volume 9011, March 2014.

- [175] A. G. Solimini. Are there side effects to watching 3D movies? a prospective crossover observational study on visually induced motion sickness. *PLOS One*, February 2013.
- [176] F. Speranza, W. J. Tam, R. Renaud, and N. Hur. Effect of disparity and motion on visual comfort of stereoscopic images. *Proceedings of the SPIE*, 6055, 2006.
- [177] G. Sun and N. Holliman. Evaluating methods for controlling depth perception in stereoscopic cinematography. In *Proc. SPIE*, volume 7237, February 2009.
- [178] K. Talmi and J. Liu. Eye and gaze tracking for visually controlled interactive stereoscopic displays. *Signal Processing: Image Communication*, 14:799–810, 1999.
- [179] W. J. Tam, F. Speranza, C. Vázquez, R. Renaud, and N. Hur. Visual comfort: stereoscopic objects moving in the horizontal and mid-sagittal planes. In *Proc. SPIE 8288, Stereoscopic Displays and Applications XXIII*, volume 8288, February 2012.
- [180] H. E. Tasli and A. A. Alatan. User assisted disparity remapping for stereo images. *Signal Processing: Image Communication*, 28(10):1374–1389, November 2013.
- [181] K. Templin, P. Didyk, K. Myszkowski, M. M. Hefeeda, H. P. Seidel, and W. Matusik. Modeling and optimizing eye vergence response to stereoscopic cuts. *ACM Trans. on Graphics*, 33(4):145:1–145:8, July 2014.
- [182] K. Terzić, L. Hotz, and J. Šochman. Interpreting structures in man-made scenes: Combining low-level and high-level structure sources. In *ICAART*, Valencia, Spain, Jan 2010.
- [183] M. Torii, Y. Okada, K. Ukai, J. S. Wolffsohn, and B. Gilmartin. Dynamic measurement of accommodative responses while viewing stereoscopic images. *Journal of Modern Optics*, 55(4-5):557–567, 2008.
- [184] C. H. Tsai, K. Lee, and K. C. Huang. Autostereoscopic display apparatus, June 2010. US Patent 7,738,171.
- [185] M. Urvoy, M. Barkowsky, and P. Le Callet. How visual fatigue and discomfort impact 3D-TV quality of experience: a comprehensive review of technological, psychophysical, and psychological factors. *Annals of Telecommunications*, 68:641–655, 2013.
- [186] M. Vinnikov and R. S. Allison. Gaze-contingent depth of field in realistic scenes: The user experience. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '14, pages 119–126, New York, 2014. ACM.
- [187] B. Wang and K. J. Ciuffreda. Depth-of-focus of the human eye in the near retinal periphery. *Vision Research*, 44(11):1115–1125, May 2004.
- [188] C. Wang and A. A. Sawchuk. Disparity manipulation for stereo images and video. *SPIE*, 6803, 2008.
- [189] F. J. Wang, X. Z. Sang, Y. D. Liu, G. Z. Shi, and D. X. Xu. Visual fatigue evaluation based on depth in 3D videos. In *Proc. SPIE 8913, Optical Storage and Display Technology*, volume 8913, August 2013.
- [190] H. M. Wang, H. C. Tseng, and K. Y. Liao. Comfortable disparity modification for existing DIBR systems. *Journal of Information Science and Engineering*, 31(5):1577–1592, September 2015.
- [191] L. Wang, K. Teunissen, Y. Tu, L. Chen, P. Zhang, T. Zhang, and I. Heynderickx. Crosstalk evaluation in stereoscopic displays. *Journal of Display Technology*, 7(4):208–214, April 2011.
- [192] X. Wang and C. Hou. Improved crosstalk reduction on multi-view 3D display by using BILS algorithm. *Journal of Applied Mathematics*, 2014, 2014.
- [193] C. Ware, C. Gobrecht, and M. A. Paton. Dynamic adjustment of stereo display parameters. *IEEE Trans. on Systems, Man and Cybernetics*, 28:56–65, January 1998.
- [194] M. A. Weissman and A. J. Woods. A simple method for measuring crosstalk in stereoscopic displays. In *Proc. SPIE 7863, Stereoscopic Displays and Applications XXII*, volume 7863, February 2011.
- [195] S. Winkler and D. Min. Stereo/multiview picture quality: Overview and recent advances. *Signal Processing: Image Communication*, 28(10):1358–1373, November 2013.
- [196] M. Wopking. Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus. *Journal of the SID*, 3(3):101–103, 1995.
- [197] C. Wu, C. T. Li, Y. C. Lai, C. C. Cheng, and L. G. Chen. Disparity remapping by nonlinear perceptual discrimination. In *3rd International Conference on 3D Systems and Applications*, Seoul, June 2011.
- [198] Z. Xia, X. Li, Y. Cui, L. Chen, and K. Teunissen. A study on the perceptual relevance of crosstalk equations for stereoscopic displays. *Journal of the Society for Information Display*, 21(6):271–279, 2013.
- [199] L. Xing, J. You, T. Ebrahimi, and A. Perkins. A perceptual quality metric for stereoscopic crosstalk perception. In *IEEE International Conference on Image Processing (ICIP)*, pages 4033–4036, September 2010.
- [200] Y. Xue, Y. Wang, L. Cao, L. Han, B. Zhou, and M. Li. The backlight control system aimed at reducing crosstalk in autostereoscopic displays. *3D Research*, 5(2), 2014.
- [201] T. Yan, R. W. H. Lau, Y. Xu, and L. Huang. Depth mapping for stereoscopic videos. *International Journal of Computer Vision*, 102(1-3):293–307, 2013.
- [202] S. Yano, S. Ide, T. Mitsuhashi, and H. Thwaites. A study of visual fatigue and visual comfort for 3D HDTV/HDTV images. *Displays*, 23:191–201, 2002.
- [203] J. W. Yoo, S. Yea, and I. K. Park. Content-driven retargeting of stereoscopic images. *IEEE Signal Processing Letters*, 20(5):519–522, May 2013.
- [204] T. K. Yu, T. W. Chan, and J. L. Tung. Method for reducing crosstalk of stereoscopic image, July 2014. US Patent App. 13/913,497.
- [205] N. Yun, Z. Feng, J. Yang, and J. Lei. The objective quality assessment of stereo image. *Neurocomputing*, 120(23):121–129, November 2013.
- [206] W. Zellinger, B. A. Moser, A. Chouikhi, F. Seitner, M. Nezveda, and M. Gelautz. Linear optimization approach for depth range adaption of stereoscopic videos. *Electronic Imaging*, 2016(5):1–6, 2016.
- [207] M. Zeng, H. Azartash, and T. Nguyen. Crosstalk modeling in circularly polarized stereoscopic LCDs. In *IEEE International Conference on Image Processing (ICIP)*, pages 3464–3468, October 2014.
- [208] M. Zeng and T. Nguyen. Crosstalk modeling, analysis, simulation and cancellation in passive-type stereoscopic LCD displays. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1840–1844, May 2013.
- [209] M. Zeng, A. E. Robinson, and T. Q. Nguyen. Modeling, prediction, and reduction of 3D crosstalk in circular polarized stereoscopic LCDs. *IEEE Trans. on Image Processing*, 24(12):5516–5530, December 2015.
- [210] F. Zeri and S. Livi. Visual discomfort while watching stereoscopic three-dimensional movies at the cinema. *Ophthalmic and Physiological Optics*, 35(3):271–282, May 2015.
- [211] L. Zhang, Y. Q. Zhang, J. S. Zhang, L. Xu, and J. B. Jonas. Visual fatigue and discomfort after stereoscopic display viewing. *Acta Ophthalmologica*, 91(2):149–153, 2013.
- [212] Y. Zhang and J. Shen. Stereoscopic 3D crosstalk prediction. In *International Conference on Image and Signal Processing (CISP)*, pages 509–513, October 2014.

- [213] J. Zhou, G. Jiang, X. Mao, M. Yu, F. Shao, Z. Peng, and Y. Zhang. Subjective quality analyses of stereoscopic images in 3dtv system. In *IEEE Visual Communication and Image Processing*, pages 1–4, Tainan, November 2011.
- [214] M. Zhou, H. Wang, W. Li, S. Jiao, T. Hong, S. Wang, X. Sun, X. Wang, J. Y. Kim, and D. Nam. A unified method for crosstalk reduction in multiview displays. *Journal of Display Technology*, 10(6):500–507, June 2014.
- [215] F. Zilly, J. Kluger, and P. Kauff. Production rules for stereo acquisition. *Proceedings of the IEEE*, 99(4):590–606, April 2011.
- [216] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. A. J. Winder, and R. Szeliski. High-quality video view interpolation using a layered representation. *ACM Trans. Graph.*, 23(3):600–608, 2004.