

Subjective assessment of display stream compression for stereoscopic imagery

Sanjida Sharmin Mohona  | Laurie M. Wilcox  | Robert S. Allison 

Center for Vision Research, York University, Toronto, Ontario, Canada

Correspondence

Sanjida Sharmin Mohona, York University, Toronto, ON, M3J1P3, Canada.
Email: sanjida.sharmin23@gmail.com

Funding information

Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: CRDPJ 523632-18; Ontario Centres of Excellence, Grant/Award Number: VIP II program, VIP contract number: 2018-0307

Abstract

High-resolution display bandwidth requirements often now exceed the capacity of display link channels necessitating compression. The goal of visually lossless compression codecs such as VESA DSC 1.2 is that viewers perceive no difference between the compressed and uncompressed images, maintaining long-standing expectations of a lossless display link. Such low impairment performance is difficult to validate as artifacts are at or below sensory threshold. We have developed a 3D version of the ISO/IEC 29170-2 flicker paradigm and used it to compare the effects of image compression in flat images presented in the plane of the screen (2D) to compression in flat images with a disparity offset from the screen (3D). We hypothesized that differences in the location and size of the compression errors between the disparate images in the 3D case would affect their visibility. The results showed that artifacts were often less visible in 3D compared to 2D viewing. These findings have practical applications with respect to codec performance targets and algorithm development for 3D movie, animation, and virtual reality content. In particular, higher compression should be attainable in stereoscopic compared to equivalent 2D images because of increased tolerance to artifacts that are binocularly unmatched or have disparity relative to the screen.

KEYWORDS

3D image quality, image compression, ISO/IEC 29170, stereoscopic display, subjective assessment, visually lossless

1 | INTRODUCTION

Ultra high-resolution stereoscopic displays in virtual reality (VR) and augmented reality (AR) are increasing in availability and popularity. With the increase in resolution, bandwidth demands for the video data interfaces increase significantly. Historically, video sources have transmitted uncompressed pixels to displays through the digital interfaces. But the bandwidth of such interfaces has not increased at the same rate as the growth in pixel bandwidth. So maintaining a high data rate across the

interfaces requires more power, more wires, and more shielding to prevent electromagnetic interference.¹ These requirements increase device weight, hardware cost, and complexity and are sometimes economically infeasible with current technology.² Bandwidth can be greatly reduced using existing algorithms such as H.264 Intra-only, Motion JPEG 2000,³ and Dirac/VC-2,⁴ but these algorithms cannot provide visually lossless quality with modest hardware complexity, in real time, and with the low latency expected of a display link.¹ To cope with these problems, the industry has developed

lightweight, very low impairment compression techniques to reduce bandwidth requirements without impacting image quality or incurring significant latency. For example, JPEG XS^{2,5} is a lightweight image compression algorithm which provides low latency for targeting production, internet protocol, and ethernet applications, but it may not deal well with specific types of content, for example, subpixel rendered text.¹ The Video Electronics Standards Association (VESA) introduced a standard for visually lossless, low cost, and interoperable image/video compression designed to work in the display link known as Display Stream Compression (DSC)⁶ and subsequently published another technique referred to as VESA Display Compression-M (VDC-M).⁷

In general, all practical image compression techniques are lossy, in the mathematical sense of being irreversible and thus losing information. In contrast to mathematically lossless coded images, *visually lossless* coded images or image sequences should be perceptually indistinguishable from the original under the same viewing conditions and for the same spatial area.⁸ However, even in the case of algorithms which aim to be visually lossless, there may occasionally be some visible differences between the original and compressed images. These differences are undesirable and distracting and may affect the subjective quality of image content. Objective and subjective measures for image assessment have been developed mostly in the context of 2D (two-dimensional) images and video.⁹ In a stereoscopic three-dimensional (S3D) display images are presented separately to the left and right eye; users perceive depth when the images are viewed simultaneously.¹⁰ These image pairs are captured from slightly different vantage points; therefore, they are not identical but contain slight offsets or disparities that are used by the visual system to interpret the relative separation of objects and points in the scene. This aspect of S3D image pairs means that coding artifacts can be common in the two images (matched artifacts) or differ qualitatively and quantitatively (unmatched artifacts). The visibility of matched and unmatched artifacts depends on how the human visual system combines and reconstructs the 3D scene from the separate images. The objectives of the present study were to

- Develop and test a novel technique for assessing visually lossless S3D compression. We modified a standard subjective image quality assessment technique designed to verify visually lossless coding performance (ISO/IEC 29170-2) for S3D display. Using this method, we perform a subjective assessment of a state-of-the-art visually lossless codec, DSC 1.2, in S3D with naive viewers.

- Compare the visibility of compression artifacts from DSC 1.2 for 2D viewing versus S3D viewing with flat images.
- Compare the visibility of compression artifacts under matched (when they are the same in both views) versus unmatched (when they differ) conditions.

We have previously presented an abbreviated description and preliminary results for the S3D assessment technique¹¹ and here extend this to describe the approach in detail and consider the effects of disparity and similarity of artifacts in the two eyes (same or different) separately.

2 | BACKGROUND

Objective image quality metrics can be used to help identify compression artifacts (e.g., flickering, aliasing, color quantization [degradation of color quality], color banding, and block boundary artifacts). Such metrics are useful for detecting issues during the development of a codec and for real-time or offline automated quality assessment. These techniques have generally been designed to quantify the expected degradation of an image relative to a reference (or with respect to expected/natural image statistics for no-reference techniques).¹²⁻¹⁴ Often, objective methods are based on models of human visual sensitivity and have been validated against subjective quality ratings to perceptually scale the impact of visible image imperfections. However, they are not intended to assess the detectability of barely visible image artifacts in complex images.¹⁵ Given that there is no perfect model of human vision¹⁶ and people vary in their sensitivity under these conditions, subjective assessment is essential. The results of subjective image quality assessment, especially at visibility threshold, may be affected by a wide range of parameters such as measurement scale, industrial versus academic assessment setting, psychometric function assumed, and the subject's task. Brunnström et al.¹⁷ identified three parameters that could affect subjective testing: display, signal (test image), and viewing distance/angle. They performed an experiment to determine the effects of type of display on detection threshold at the same EOTFs (Electrical to Optical Transfer Functions) and viewing conditions. They found that contrast detection thresholds differed reliably for the two images from the Categorical Subjective Image Quality (CSIQ) masking database¹⁸ that they used, but that display type only had a significant impact for one of the three subjects.

Hoffman and Stolitzka¹⁵ introduced a new subjective method for assessing whether a compressed image was visually lossless that was subsequently adopted as

ISO/IEC international standard 29170-2.⁸ The procedure was based on a two-alternative forced choice detection task often used to estimate sensory detection thresholds. In a large-scale study, Allison et al.¹⁹ used this standard to conduct a subjective evaluation of DSC 1.2 for 2D Standard Dynamic Range (SDR) images at different levels of compression, different chroma subsampling, and different slice sizes. They implemented two different protocols described in the ISO/IEC 29170-2⁸ standard: Flicker and Panning. The only difference between the two protocols was that in the flicker protocol the test image sequence consisted of the compressed image alternating with the uncompressed image, whereas in the panning protocol the images were moved diagonally, and the test sequence was compressed whereas the reference sequence was not. The user's task was to identify which image looked worse or had flicker or artifacts. They found that DSC 1.2 showed visually lossless performance for RGB and YUV subsampled sources at target levels of compression (down to 8 bit per pixel [bpp]) although some challenging images were not visually lossless.¹⁹ Their study also showed that in most cases, for moving content in the panning protocol, viewers were less sensitive to compression artifacts relative to the flicker protocol. This sensitivity could be due to motion silencing (failure to detect a change in image properties due to the presence of coherent motion).²⁰ Choi et al.²¹ demonstrated a strong motion silencing effect in subjective tests of detection of flicker distortion in naturalistic videos. They found that the speed of object motion affected the visibility of flicker distortion. For objects with fast motion, subjects perceived less flicker distortion despite attending to and tracking the moving object. Motion silencing becomes more apparent in poor quality video.²¹

Most of the previous empirical and theoretical evaluation of image compression assessment has been performed using 2D images or video presented on 2D displays. S3D image compression introduces additional considerations including the impact of artifacts on depth perception and the effects of matching of the compression artifacts in the two eyes. Chen et al.^{22,23} studied human detection of local distortion in S3D images. They assessed perception of distortion by varying severity and type of distortion for low-level image content to determine the effects of masking on S3D images. They found that in S3D-images, image contrast was significantly correlated with the visibility of blur, JPEG, and JP2K artifacts, whereas range energy (range of depths in the captured scene) was correlated with the visibility of blur and JP2K artifacts. For blur and JP2K distortions, regions with higher contrast or range energy were susceptible to more visible artifacts. On the other hand, contrast and

range variation did not have significant effects on white noise visibility ($p > 0.05$). Although these authors (among others²⁴⁻²⁶) have assessed binocular effects of noticeable image distortion, it is not clear if the results apply to near visually lossless compression where artifacts appear near their detection threshold. In this experiment, we compressed images using a codec targeting visually lossless compression and determined whether near-threshold binocular or dichoptic artifacts are suppressed or enhanced by S3D presentation.

3 | RATIONALE AND AIMS

Although objective and subjective measures for S3D image assessment have been developed,²⁷⁻³⁰ to our knowledge no techniques for assessment of visually lossless S3D images (i.e., at or below detection threshold) have been developed or assessed. We have presented preliminary results for S3D images using our modified flicker paradigm¹¹; here we formally describe, assess and implement this novel protocol for subjective assessment of visually lossless quality of S3D images. Under controlled conditions, we apply this protocol to the assessment of a state-of-the-art codec targeting visually lossless compression (DSC 1.2 with 3:1 compression). Although this protocol can be used to assess quality in any type of S3D imagery, in this study, disparity was deliberately limited to isolate the effects of offset from the screen (disparity pedestal) and binocularly unmatched noise rather than to study the effects of depth variations in the scene. To do so, we used 2D reference images to derive the left and right test images and assessed the effects of (1) disparity of the images relative to the screen and (2) the matching of the compression errors in the left and right images. The use of 2D images with a disparity offset (flat stereoscopic 3D images) allowed us to isolate the effects of disparity. That is, we can measure the effects of displaying the images with a constant disparity with respect to the screen (a disparity pedestal) without the complications of variation in disparity within the image as would be present in true S3D images. Further, the use of 2D images also allowed us to assess the effects of compression artifacts when they were identical in the two eyes' images compared to when they differed. By using identical 2D test images in the left and right eyes, the image compression in the two half images could be matched perfectly. Alternatively, by offsetting the images before compression, different compression errors could be introduced in the left and right images. Thus, in a controlled fashion we evaluated the impact of these two important factors (disparity and compression matching) on achieving truly visually lossless S3D compression.

4 | METHODS

4.1 | Observers

Sixteen subjects were recruited, but after visual screening four were excluded (two for poor stereoscopic vision and two for poor visual acuity [see visual screening criteria below]), and one subject was excluded for failing to detect flicker in catch trials. Among the remaining 11 subjects (18 to 29 years), 8 were female and 3 were male. All the participants were compensated for their time either financially or through course credit. All observers gave informed consent and the protocol was approved by the York University research ethics board.

4.2 | Apparatus

The apparatus consisted of two HP Dreamcolor Z24x monitors arranged in a mirror stereoscope configuration (Figure 1), mirrors were placed at 90° to each other, $\pm 45^\circ$ angle to the participants' face (frontal plane). The screen size was 51.5 cm (59.6°) \times 32.5 cm (39.7°), the pixel resolution was 1920 \times 1200 at 60 Hz. The monitors were carefully matched for color and luminance (maximum 119 cd/m² and minimum 0.14 cd/m²). The viewing distance was 45 cm (30 pixels per degree) and participants used a gamepad (Microsoft SideWinder Plug and Play Game Pad) to make their responses.

Prior to testing, the Snellen Chart, Randot stereo test (2015 Stereo Optical Company, Inc.) and Ishihara test (Kanehara Trading Inc., 24 plate edition, 2005) were used to measure visual acuity, stereoacuity, and color vision, respectively. For inclusion, participants had to correctly identify all the color test plates, discriminate disparity of at least 40 seconds of arc, and have 20/20 corrected vision. The Psychtoolbox Package with MATLAB

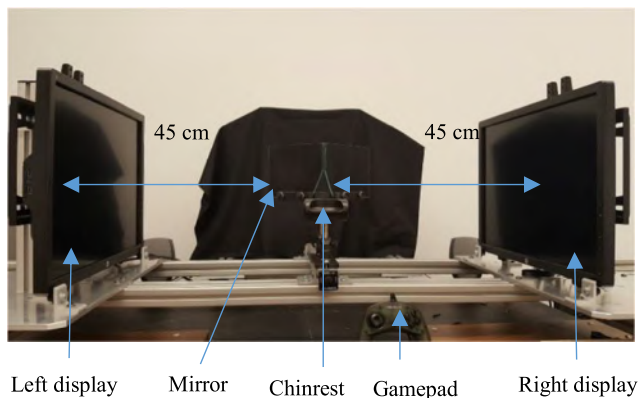


FIGURE 1 Test station

R2016a were used to develop the script to display the stimuli and to record participants' responses.

4.3 | Stimuli

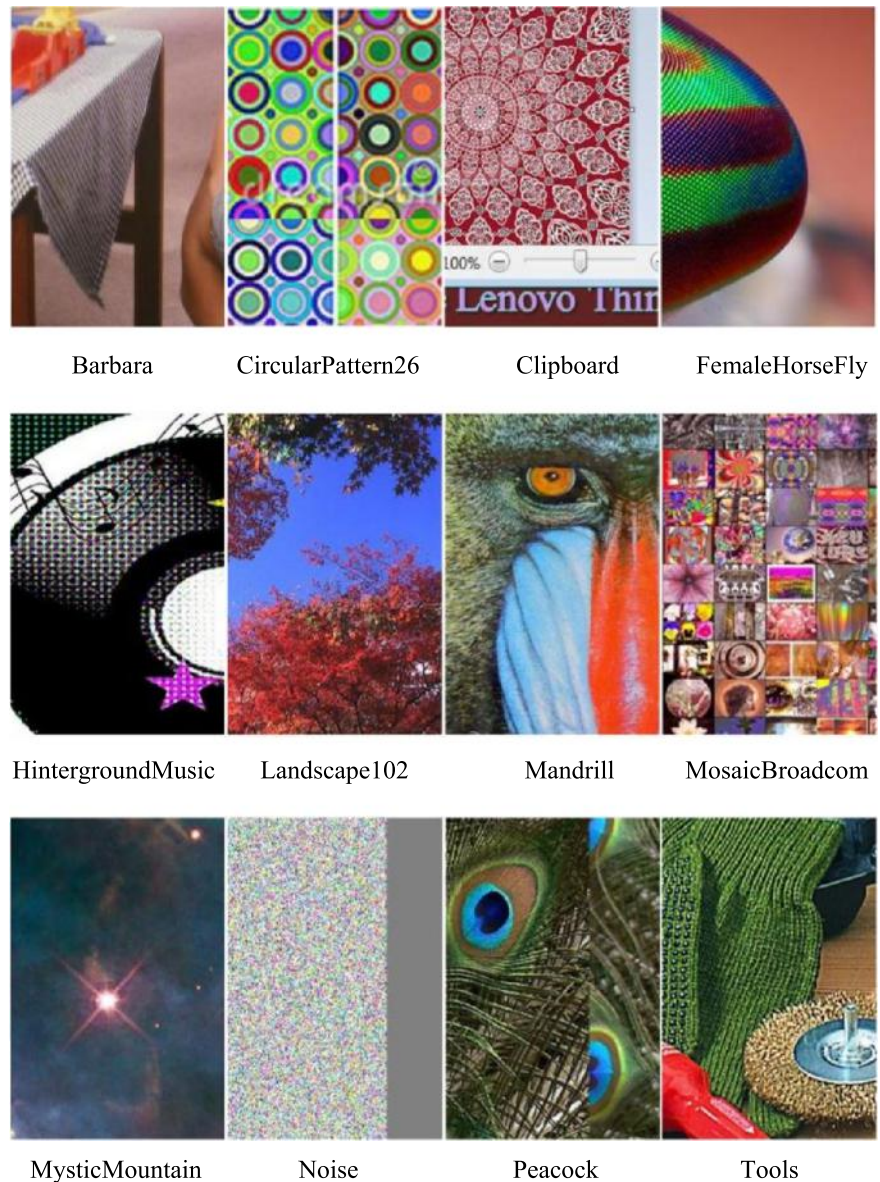
Twelve 2D images (see thumbnails in Figure 2) were used for this experiment. Images were selected so that they contained a variety of content including: animals, text, people, noise, and computer graphics. For all images a 200 \times 300 pixel crop location was chosen to constrain observers' attention to the region of interest.

DSC 1.2 compression was applied to full-size images which were then cropped. The original images were 24-bit full-color standard dynamic range RGB images (8 bits per pixel per color channel—bpc, 24 bit per pixel—bpp) that were compressed 3:1 using DSC 1.2 (slices per line = 1) to produce 8 bpp images. For display, these images were decompressed to produce 24-bit RGB images containing any artifacts introduced by the compression step. We only used a single compression level as the main comparisons were between similarly compressed image pairs (i) with or without a disparity pedestal and (ii) matched or unmatched compression artifacts. These images were chosen to challenge the compression codec at the target level of compression and as a result codec effectiveness varied across the images.

4.4 | Task

The ISO/IEC 29170-2 (Annex B)⁸ flicker procedure (Figure 3) was modified for using with stereoscopic displays. In the stereoscopic version of the flicker test, both the test and reference stimuli are stereoscopic image pairs consisting of a left and right half-image that form a stereoscopic image when combined by the brain when viewed in the stereoscopic display. The test sequence consisted of the compressed stereoscopic image temporally interleaved (alternating) with the uncompressed version at a fixed frequency. In the reference sequence, the uncompressed S3D image alternated with itself. When the images were interleaved, the uncompressed–uncompressed reference sequence appeared as static, whereas, if artifacts were present, the compressed–uncompressed test sequence could contain regions of flicker due to the temporal sensitivity of the human visual system. Figure 3 shows an illustration of the 3D flicker test protocol. The flicker rate was 5 Hz (100 ms for presentation of each phase), as recommended in the ISO/IEC 29170-2 protocol. Participants saw the test and reference sequences side by side; on each two-alternative

FIGURE 2 Thumbnails of the 12 test images



forced-choice trial the participant was asked to identify the compressed image (i.e., which image sequence contained flicker). As specified in the ISO/IEC protocol, catch trials were introduced using easily detectable control images. These control images were highly distorted using the JPEG 2000 codec with compression quality 10. Participants' data were only included in the final analysis if they detected the control images on more than 95% of the trials.

4.5 | Design

The study employed a 2×2 within-subjects design. There were two independent variables each with two levels: Depth of images (no disparity, disparity) and

Compression (same, different). The difference in image compression was produced by shifting the image location prior to DSC compression. Even when the identical codec and compression levels are applied, the lateral shift produces small stochastic difference between compressed images. The dependent variable was the detection rate of the compressed image (average proportion of correct detections). Each of the 12 images in each condition was presented 20 times resulting in 240 trials per condition (1200 trials for five conditions) and each catch trial image was presented 20 times in each session resulting in 120 catch trials per session (total 240 catch trials in 2 sessions). The total number of trials per observer was 1440 including the catch trials and viewing time for each trial was 4 s.

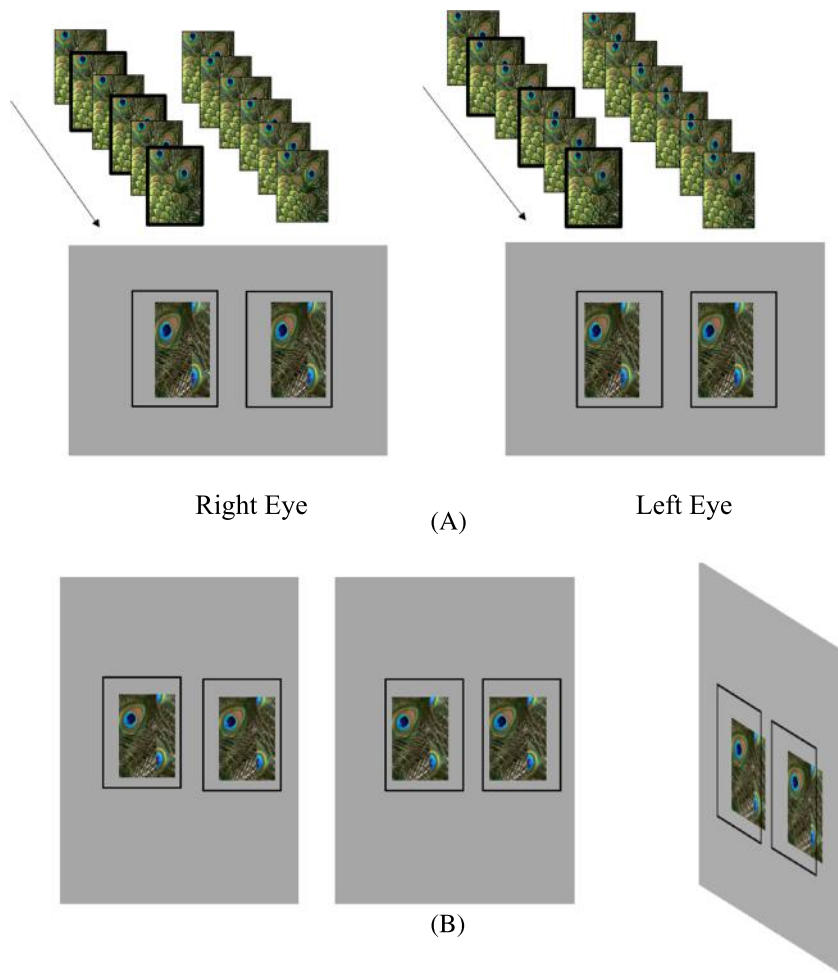


FIGURE 3 Illustration of the 3D flicker protocol. (A) The upper part depicts the temporal image sequence. One half of each of the left and right eye display consisted of the reference image alternated with the test image (indicated by the bold outline). This test image is shown on the left in this example but was randomly presented on either the left or the right side of the combined S3D display. The other half of each eye's display consisted of the reference image alternating with itself and so appeared as a static presentation of the reference image. The bottom left and right sides of (A) show a representation of the view presented to the left and right eyes, respectively (if fused by crossing the eyes, the images will appear displaced in depth behind the reference frame). (B) In the disparity conditions the images would appear in stereoscopic depth behind the screen when the two eyes' images were combined

4.6 | Display conditions

The stimuli were compressed and viewed under four display conditions:

1. Disparity, with the same compression in the left and right images.
2. Disparity, with different compression in the left and right images.
3. No disparity, with same compression in the left eye and right eye image with both images corresponding to
 - a. the left half image from condition 2.
 - b. the right half image from condition 2.
4. No disparity, with different compression in the left and right images.

All conditions were dichoptic and condition 2 corresponded to the compression typically used when displaying S3D. As outlined previously, we used a selected (i.e., cropped) portion of the image instead of a full image. The disparity used for conditions 1 and 2 was 10 pixels ($1/3^\circ$ visual angle). Therefore, to prepare test images with disparity we introduced a 5-pixel horizontal

shift in opposite directions in each of the 2D images for a total of 10 pixels offset. This shift was applied either after or before image compression to create the same compression and different compression cases, respectively. Figure 4 presents the image processing steps for each condition. Introducing the image shift prior to compression (condition 2) caused the images to differ and produced small stochastic variations in the compressed images even at the same compression levels.

For condition 1, shifting was performed after compression to ensure the compression was the same in both eyes' images. The full-size compressed image was cropped for the left eye image. Then, to introduce disparity into the stereoscopic pair presented on the displays, the full compressed image was shifted (but the crop position was not shifted) to produce the right eye image. To produce the reference (i.e., uncompressed) image pair for condition 1, the original image was cropped to produce the left eye image and then to produce the right eye image, the original image was shifted and cropped in the same position as for the left eye image. For condition 2, the original image was compressed and cropped to produce left eye image; for the right eye image, the original image was

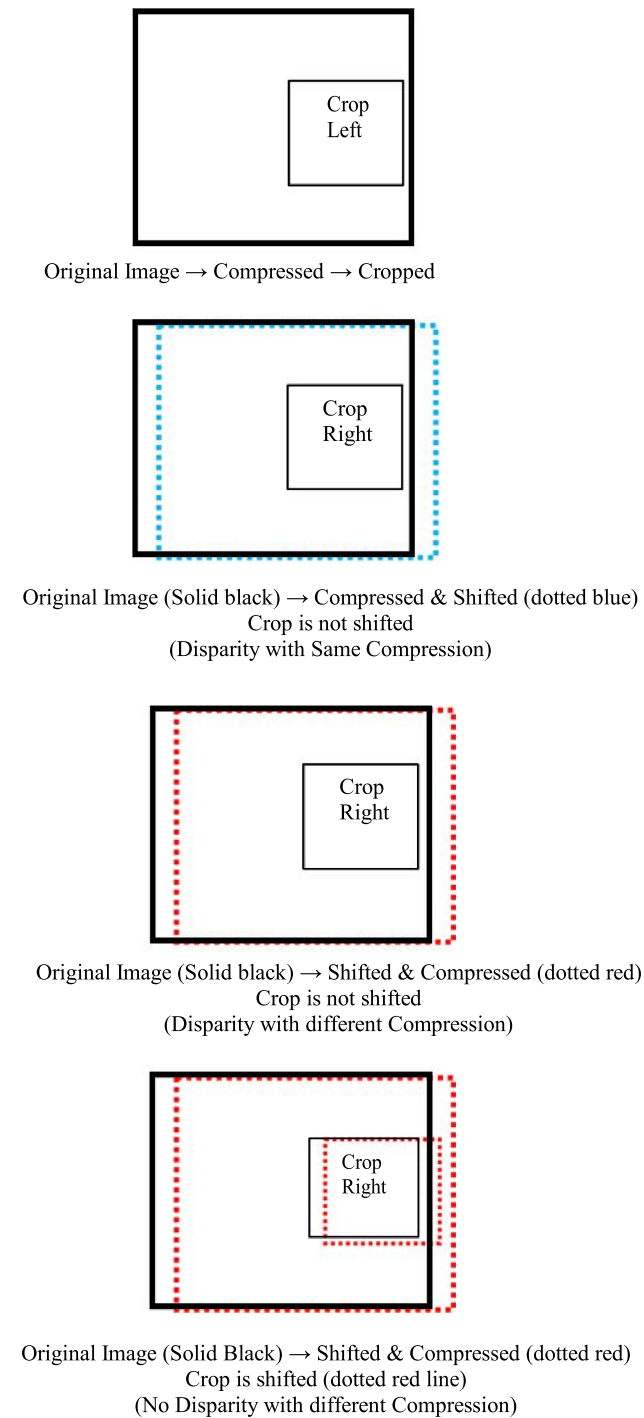


FIGURE 4 Illustration of the four dichoptic viewing conditions. The solid black outline represents the original image. Crop regions are shown as smaller rectangles. Dashed rectangles indicate position of original image following shifting either preceded by (blue) or followed by (red) compression. In the first three panels the crop region is not shifted, in the bottom panel the crop region is shifted along with the original

shifted before compression, the image was compressed and after compression the image was cropped in the same position as used for left eye image. Thus, the displayed

image pair has disparity and different compression in the two eyes. To prepare the reference image pair for condition 2, the original image was cropped to produce the left eye image and the right eye image was prepared by shifting original image and then cropping in the same position as for left eye image (the procedure and thus the reference produced was identical to that in condition 1). For condition 3 either the left (compressed and cropped) or right (compressed and cropped) image was displayed to both eyes so the displayed image had no disparity and the same compression to both eyes. Here, the compressed left image was obtained by cropping the original image after compression and the reference left image was obtained by cropping the original image without compression. To create the right eye reference image, the original image was cropped after shifting 10 pixels horizontally.

To obtain the compressed right image, the original image was shifted 10 pixels, then compressed and cropped. Both the compressed left and compressed right images were used in case there was a possible difference in artifact visibility in the two cases due to the image shifting. For condition 4, the original image was compressed and cropped to give the left eye test image; then for the right eye test image the original was shifted and then compressed so images had different compression but the crop region was shifted by the same amount as the image shift, resulting in zero disparity. The left eye and right eye images for the reference pair were the same image, which was produced by cropping the original image before compression at the same location. To show the impact of compression on the images in condition 4, in Table 1 we provide the PSNR between left eye and right eye image and between compressed versus reference image.

4.7 | Procedure

The experiment was performed by each participant in two 60-min sessions over two separate days. To limit viewer fatigue, each session was divided into 5–6 min blocks between which participants could take rest breaks.* After arrival, participants were briefed about the goals of the research and informed consent was requested. They were then screened using the three vision tests and age and gender were recorded. Before starting the experiment, verbal instructions were given to each participant and they participated in a brief practice session with two control images (two trials per image). Following practice, the main experiment was started. The task was to detect the flickered image in each trial. Observers could respond at any time during the 4-s trial

TABLE 1 PSNR between left eye versus eye image, and compressed versus uncompressed image

Images	Compressed L versus reference L of condition 4	Compressed R versus reference R of condition 4	Compressed L versus compressed R of condition 4
Barbara	42.7223	42.824	40.55176
CircularPattern26	34.25251	34.24961	31.61995
Clipboard	36.72344	36.57763	34.20463
FemaleHorseFly	35.83467	35.87543	33.38604
HintergroundMusic	34.96581	35.06939	32.48409
Landscape102	31.49795	31.54009	29.06393
Mandrill	33.92622	33.89638	31.36713
MosaicBroadcom	34.76721	34.79389	32.23884
MysticMountain	40.73737	40.74989	38.32638
Noise	25.92884	25.97993	24.10274
Peacock	38.11989	38.14685	35.7263
Tools	31.60354	31.61182	29.17022

and feedback (a tone) was provided when they failed to select the target image (i.e., the compressed image). If they did not respond within 4 s, a blank screen appeared with text asking the participant to provide their response.

5 | RESULTS

The primary hypothesis is that the sensitivity to compression artifacts would differ between conditions with and without disparity, that is, where the stimulus appeared at the screen plane (Hypothesis 1). The second main hypothesis was that sensitivity to compression artifacts would differ when the compression was matched (same) versus unmatched (different) in the two eyes (Hypothesis 2).

To test the two hypotheses, the flicker paradigm data was fitted using a Generalized Linear Mixed Model (GLMM)³² following the approach described by Cutone et al.³³ for the analysis of 2D image quality assessment data. GLMM analysis was performed using the R statistical software environment (R Core Team 2017) and maximum likelihood estimation was used for fitting data using lme4's "glmer." Before fitting GLMM, control data were excluded because these are unrelated to the experimental hypotheses and including control data can cause convergence failure as the correct response proportions were typically 1.0 for control data. To assess the effects of depth, compression, and their possible interaction with artifact visibility the following model was used:

$$\text{correct} \sim \text{depth} * \text{compression} + (1|\text{subject}) + (1|\text{ref}) \quad (1)$$

TABLE 2 Interaction between depth and compression

	χ^2	DF	Pr(> χ^2)
(Intercept)	10.8171	1	0.001 **
Depth	16.6773	1	4.4e-05 ***
Compression	14.7820	1	0.0001***
Depth:compression	6.4069	1	0.01 *

Note: variables: 2 depth conditions (disparity, no disparity), 2 compression conditions (same, different), 11 subjects, and 12 reference images.

*** $p < 0.001$. ** $p < 0.01$. * $p < 0.05$.

The model is expressed using Wilkinson and Rogers³⁴ notation in which the measured response variable is on the left-hand side of the “~” operator and the right-hand side describes the model in terms of the predictor variables or factors. For the above formula, the model treats depth and compression as fixed effects. Both subject ID and image were modeled as random factors which is indicated by the grouping notation (one subject) which indicates that a random intercept model is used for the within subject variable. The error distribution of the response variable was modeled as binomial as the response was dichotomous (left or right). The ANOVA for the above formula is in Table 2.

The result of GLMM analysis showed that, across all subjects and images, (1) the main effect of depth, compression and (2) the interaction of depth and compression on the perception of compression artifacts were significant. On average, artifacts are less visible when compression is the different in the two eyes than for the same compression (Figure 5) but this is particularly true for images with disparity which explains the interaction. It's important to understand the nature of this interaction

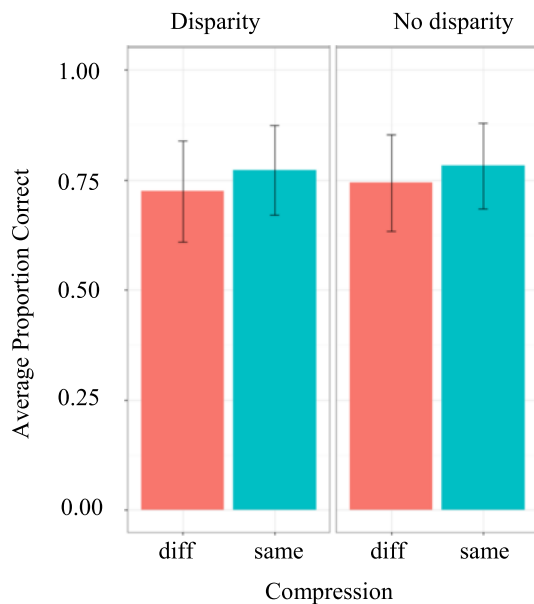


FIGURE 5 Results are shown here for different compression versus same compression for disparity and no-disparity conditions averaged across observers and images. Error bars represent ± 1 standard error of the mean (SEM)

before addressing our main hypotheses. To do so, we explored the interaction effect for each image and found that it was limited to two specific images of the 12 tested, Mandrill and Peacock. For these two images only, the depth-compression interaction had a significant effect on artifact perception. Thus, while we need to be mindful of this interaction during the hypothesis testing because it is isolated to a small subset of the images, we can consider our main hypotheses.

To test Hypothesis 1, a series of planned comparisons was performed between the proportion correct for the no-disparity and disparity conditions for each image (marginal means across compression level).

This analysis was based on tests of linear contrasts estimating the difference between no-disparity and disparity predictions for each image using the lsmeans package in R to obtain the least square mean predictions. The two-sided pairwise comparisons corresponding to Hypothesis 1 are visualized in Figure 6. False Detection Rate (FDR) p value correction was applied for the tests of these hypotheses at a significance level of 0.05. From Figure 6 we can see that, among the five images with significant differences (CircularPattern26, Clipboard, Landscape102, Mandrill, Tools), artifacts were more visible in the no-disparity condition than in the Disparity Condition for all images except Clipboard. Thus, for these images overall DSC 1.2 was more visually lossless in the disparity condition than the no-disparity condition.

To test Hypothesis 2, we compared different compression versus same compression conditions. Figure 7 shows that there was a significant difference between Different and Same Compression with DSC 1.2 for Clipboard, FemaleHorseFly, Landscape 102, MosaicBroadcom, Noise and Tools. Except for Clipboard, in all cases artifacts were less visible with different compression than with same compression.

6 | DISCUSSION

In this experiment, we assessed the effects of (1) disparity of the images relative to the screen and (2) the similarity of the compression errors in the left and right images on artifact detection. Generally, we found that when there was a difference in the visibility of compression artifacts they were less apparent in (1) 2D images presented with a disparity (“disparity” condition, offset from the screen) than in 2D images presented at the screen plane (“no-disparity” condition) and (2) when the compression errors in the two eyes differed compared to when they matched. The effect of similarity seen here is consistent with the conclusions of Chen et al.²² who reported that the conspicuity of many image artifacts increased with range energy which produces increasing differences between the images. However, we also found that increasing disparity with a depth pedestal decreased the likelihood that the compressed image would be detected, despite increasing the range of depths. Thus it appears that the differences between images and particularly differences in compression artifacts, rather than stereoscopic depth, underlies the range effects reported in Chen et al.²² The images used here, and their specific crop regions, were previously used for extensive 2D testing of the DSC protocol.¹⁹ Our 2D test results are consistent with those reported by Allison et al.¹⁹ thus confirming the generalizability of the results. In this paper we evaluated subjective performance; the interested reader can refer to other works^{15,35,36} for comparison of objective metrics responses to subjective responses for these images under 2D viewing.

6.1 | Effects of a disparity pedestal

The use of 2D images offset in depth allowed us to assess the effects of displaying the images with a constant disparity with respect to the screen (a disparity pedestal) without the complications of variation in disparity within the image as would be present in true S3D images† as assessed in our preliminary study.^{10,37} In the current experiment, with one exception, for images where the disparity and no-disparity performance of the codec

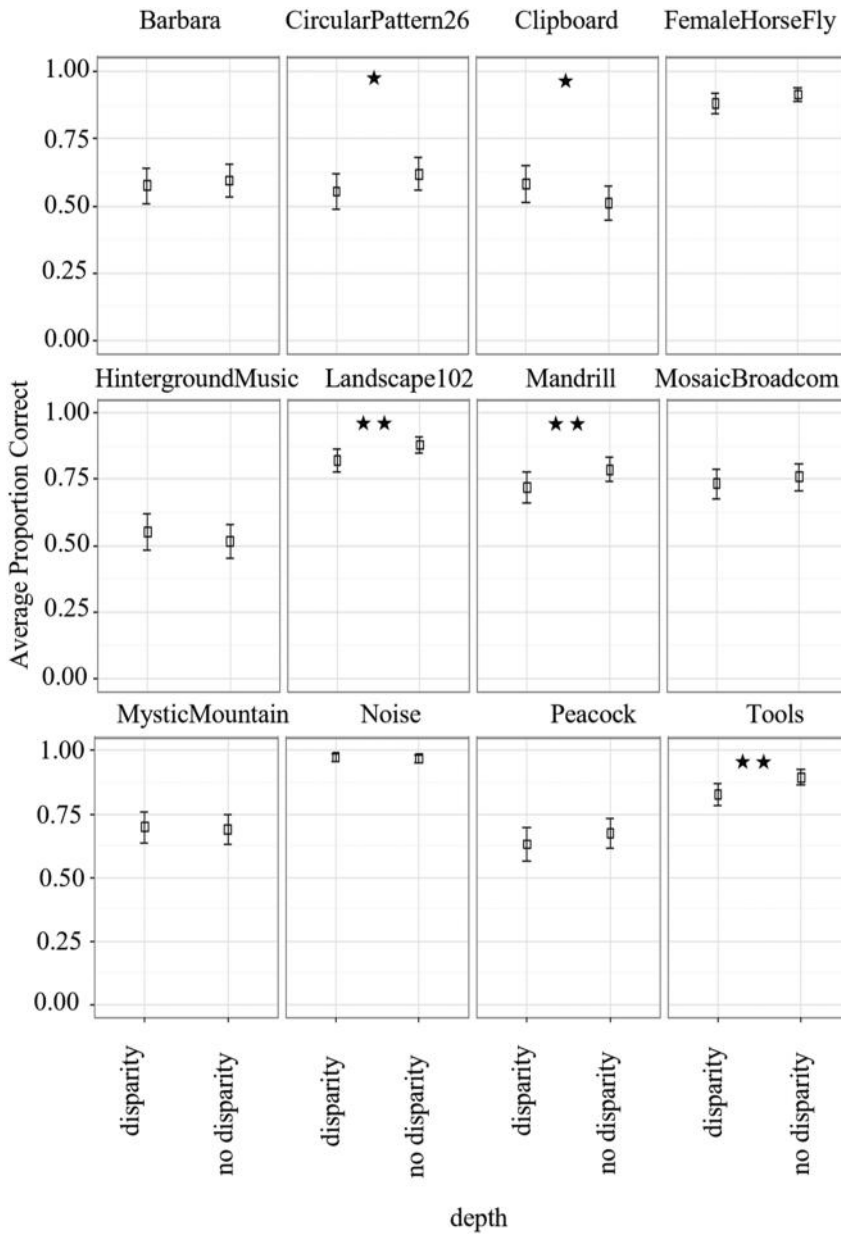
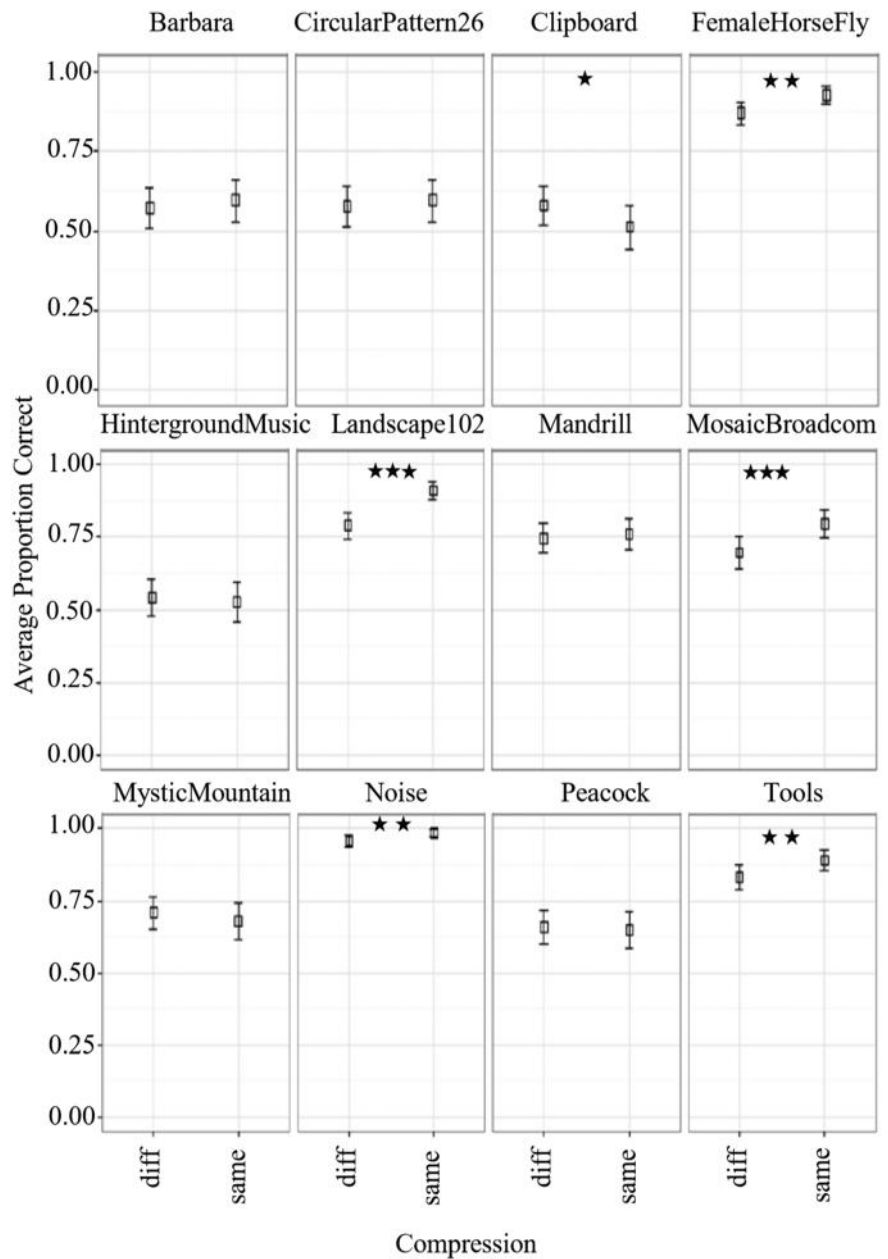


FIGURE 6 Average proportion correct for each image for the disparity versus no-disparity conditions (data for same and different compression are collapsed). Error bars represents ± 1 SE (significance: ★ ★ ★ $p < 0.001$; ★ ★ $p < 0.01$; ★ $p < 0.05$)

differed, the artifacts were more visible when the images were presented at the screen plane (no-disparity condition) than with a disparity offset (Disparity Condition). This disparity pedestal effect cannot be due to stereoscopic factors such as spurious disparity or monocular features, distortion, rivalry and so on as the images in the two eyes were identical in the same compression case, except for an offset. Many perceptual judgments are most precise when the stimulus is foveated by both eyes and performance degrades with a disparity pedestal or fixation disparity.^{38,39} However, this is unlikely to be the explanation for the pedestal effect in this experiment as the subject's fixation was not controlled and given the long viewing time, there was ample opportunity to converge on the stimulus during testing.

Stereoscopically, the no-disparity condition images were presented at the plane of the display, whereas the Disparity Condition images were behind the screen plane. However, in both conditions optimal focus of the eyes was at the screen distance. In the natural environment, vergence and accommodation are coupled, therefore when converging at a further distance the viewer should also accommodate at the further distance. If converging on the disparate stimulus caused the observer's accommodation to shift away from the screen plane,⁴⁰ there would be more defocus blur in the image with a disparity pedestal than in the images presented on the screen plane.^{41,42} However, the depth offset used in this experiment was modest and the predicted blur if the eyes focused on the far target would only be about

FIGURE 7 Average proportion correct for each image for the different (unmatched) compression versus same (matched) compression conditions (data for disparity and no disparity are collapsed). Error bars represent ± 1 SE (significance: ★★★ $p < 0.001$; ★★ $p < 0.01$; ★ $p < 0.05$)



0.1 D, within the normal depth of field of the eye,⁴³ so any blur effect would be small and likely below detection threshold. A modest mismatch between vergence and accommodation could also increase the difficulty of fusing S3D images and potentially introduce errors in fixation, or fixation disparity.⁴⁴ This might also have reduced the users' ability to detect artifacts in disparate images compared to the same images presented at the screen depth.

6.2 | Effects of matched compression

The use of 2D source images also allowed us to assess the effects of the matching of compression artifacts. We

anticipated that artifacts would be more detectable in the same compression condition as binocular summation would accentuate artifacts that were common in both eyes. Due to binocular summation the detection threshold for a stimulus is lower with two eyes than one eye.⁴⁵ Visual acuity, contrast sensitivity, flicker perception, and brightness perception⁴⁶ can be improved by binocular summation. Consistent with predictions of binocular summation, generally, DSC 1.2 was more visually lossless when unmatched compression was applied to the two eye's images compared to when matched compression was applied. An alternative explanation is that dichoptic masking could underlie the reduced perception of artifacts in the different compression condition. Dichoptic masking occurs when the two eyes view

different but similar images⁴⁷; the detection of a test stimulus in one eye is reduced by a “mask” stimulus in the other eye.

Although this increased sensitivity to matched compared to unmatched artifacts was the general finding, for two specific images among the 12 images, Mandrill and Peacock, the depth-compression interaction had a significant effect on artifact perception. For both images, artifacts were more visible with same compression than different compression when the images had disparity (consistent with our general findings) but in the no-disparity condition artifacts were more visible with different compression. The zero-disparity same compression case consisted of two sub-conditions, one where the compressed image corresponded to the left image of the zero-disparity different compression case (condition 3a) and the other corresponded to the right image of the zero-disparity different compression case (condition 3b). In most cases the performance on these two cases was very similar but for the Mandrill and Peacock images the artifacts were noticeably more visible in the left image. The Mandrill and Peacock images both contained very fine texture a factor that may be responsible for the significant differences in artifact detection in conditions 3a and 3b.

Figures 8 and 9 highlight the locations in the images (with disparity) where the difference between different compression and same compression were visible. These figures were obtained using S-CIELAB⁴⁸ to estimate the locations of artifacts in both the left and right compressed images. This produced two errorImages (difference

between compressed and reference images) calculated in CIELAB delta E units which were thresholded to highlight errors that are 4 delta E units or larger.

To produce Figures 8 and 9, these error images were combined as separate color channels to highlight differences in the location of artifacts. Because these images have disparity, the common artifacts in the left and right images appear twice, with an offset corresponding to the disparity (so all artifacts are repeated except at the edges in the same compression case). We can see that in the different compression case that for the indicated locations (enclosed in boxes on the errorImages (Figures 8A,B and 9A,B) and shown by arrows on the source image (Figures 8C and 9C)) the artifacts are not repeated or are attenuated in one copy indicating more artifacts (noise) in one image than in the other. For these two images the artifacts are more apparent in the left than the right eye consistent with the difference in detection of these artifacts in conditions 3a and 3b.

6.3 | Possibly enhanced stereoscopic artifacts

In the results described above the general pattern is for disparity to reduce the detectability of artifacts, only in the Clipboard image did we see a different outcome. In this case, artifacts were *more* perceptible for different compression than same compression and more visible in the disparity pedestal condition than in the no-disparity condition.

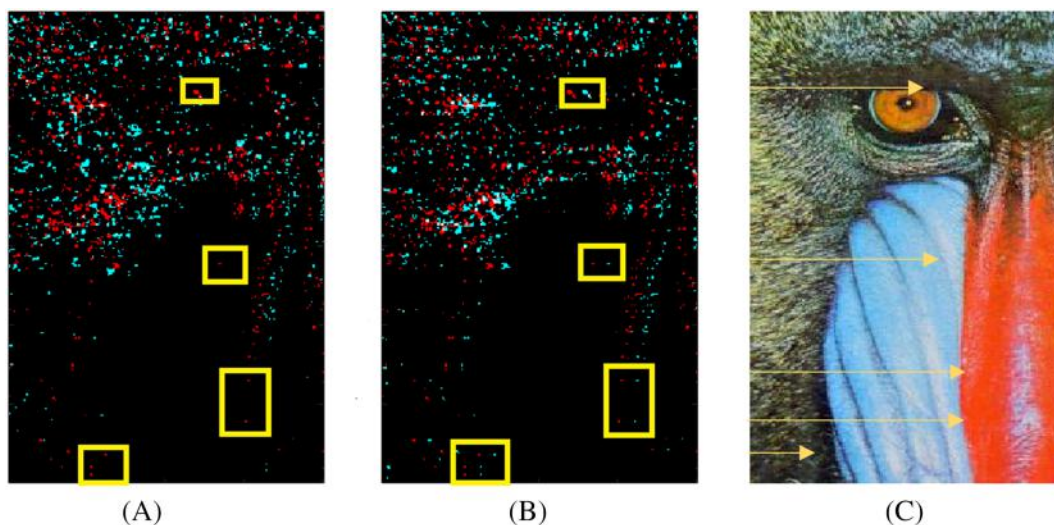


FIGURE 8 Spatial distribution of the compression errors between left and right images for Mandrill with (A) different compression versus (B) same compression, both for the disparity condition. The error images were estimated using S-CIELAB and thresholded to show locations with $\Delta E_s \geq 4$ (red indicates error in left eye, and cyan indicates error in right eye). The boxes indicate locations where artifacts in one image are not repeated or are attenuated in the other. (C) The original Mandrill image (left eye) with arrows indicating approximate locations of the boxes in the error images. See text for details

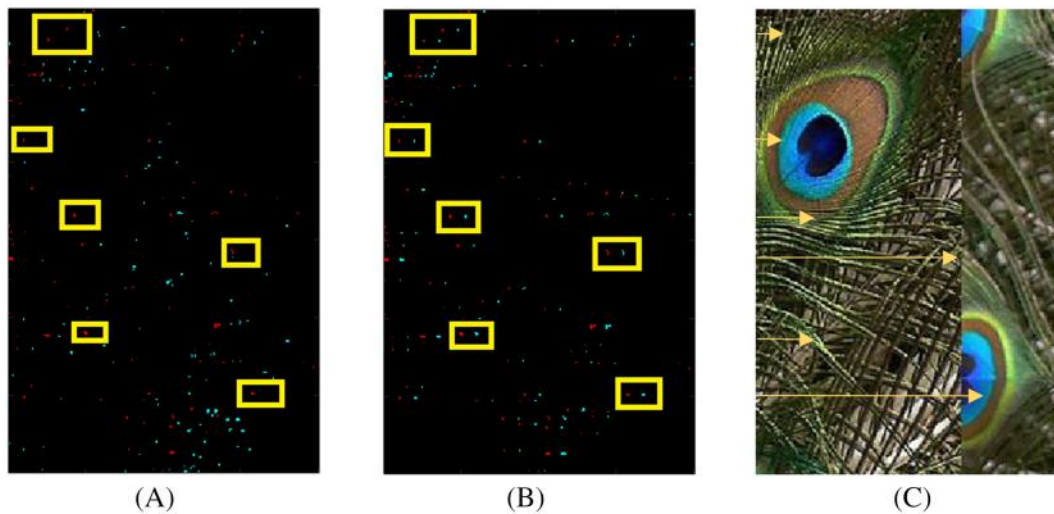


FIGURE 9 Spatial distribution of the compression errors between left and right images for Peacock with (A) different compression versus (B) same compression, both for the disparity condition. The error images were estimated using S-CIELAB and thresholded to show locations with $\Delta E_s \geq 4$ (red indicates error in left eye, and cyan indicates error in right eye). The boxes indicate locations where artifacts in one image are not repeated or are attenuated in the other. (C) The original Peacock image (left eye) with arrows indicating approximate locations of the boxes in the error images. See text for details

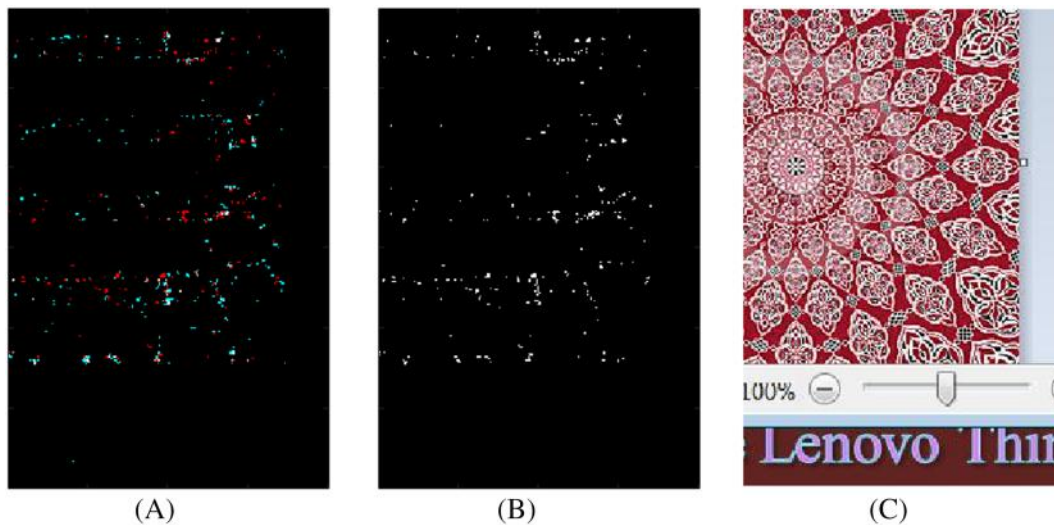


FIGURE 10 Spatial distribution of the compression errors between left and right images for Clipboard, with (A) different compression versus (B) same compression, both for the no-disparity condition. The error images were estimated using S-CIELAB and thresholded to show locations with $\Delta E_s \geq 4$ (red indicates error in left eye, and cyan indicates error in right eye; white areas have errors in the same position in both images). (C) the original clipboard image (left eye). See text for details

The compression noise generated in the two eyes' images of Clipboard may be uncorrelated which could make the artifacts more apparent in the disparity condition compared to the no-disparity condition. Figure 10 shows differences in compression artifacts between the left and right Clipboard images for the case where the images are in correspondence (with zero disparity). Thus, in the same compression case when the left and right thresholded error images are combined, all the artifacts appear as white (Figure 10B) because the images are

identical and thus both eye's errors are in the same location. In contrast, for the different compression case there are numerous artifacts that appear in one eye but not the other (Figure 10A).

Note that the artifact detection rate for Clipboard is near chance. Similarly, several of the images that showed non-significant effects of disparity or compression type had performance near chance (Barbara, CircularPattern and HintergrundMusik). Conversely, performance for Noise was near ceiling. Caution should be taken when

making comparisons between conditions for such cases because of the low statistical power.^{49,50}

The compression error images show locations where the artifacts do not match in the two images (shown as red or cyan, matching error appears white). It can be seen that there are significant regions where the artifacts do not match in the two eyes. The most obvious difference between Clipboard and other images is that Clipboard contains text (Figure 2). However, there appears to be little compression noise near the text suggesting this is not the cause of the unusual pattern of results with this image. Instead it seems more likely that the strength of the difference signal in the different compression case (i.e., most of the artifacts do not match) is responsible for the high artifact visibility for this image. We noted that monocular artifacts are introduced along the edge of the image when shifted, this may underlie the increased detection rate in the disparity case.

6.4 | Implications for stereoscopic displays

Our results suggest that in general, barely detectable compression artifacts should be less problematic in stereoscopic 3D displays compared to non-stereoscopic displays at similar compression levels. This increased tolerance to compression noise results from two factors that should limit the visibility of distortions in stereoscopic displays. First, we have shown an increased tolerance for binocularly unmatched, near-threshold artifacts in the two half images compared to common or matched artifacts. Previous work^{22,23} found similar binocular suppression of suprathreshold distortions. As a result, if all artifacts were common in the two half images of an S3D display we would expect similar performance for this display compared to equivalent 2D compressed images (viewed bi-ocularly). If the left and right images were independently compressed and some of the artifacts became unmatched in the two eyes then we would expect less sensitivity to image compression in the S3D case. Thus, on this basis we predict that in general S3D displays are at least as immune, and likely more immune, to distortion than 2D displays. The second factor is that when a flat image is presented on a disparity pedestal sensitivity to distortion decreases compared to viewing the same image on the screen plane. As a true S3D image has a variation of disparity within the image, we expect at least parts of the image to be disparate with respect to the screen and this should further increase tolerance to distortion in the S3D compared to 2D case. Indeed, in a recent conference paper we have provided data that

confirm that in most cases S3D imagery is less sensitive to compression distortions than the equivalent 2D imagery.³⁷

Our results also have implications for different modes of stereoscopic presentation and types of binocular content. For example, we deliberately used dichoptic presentations of planar images that were nearly identical for reasons of experimental control and manipulation. These are not typical S3D content, but such dichoptic images correspond to important niche use cases of stereoscopic displays. An infinitely distant scene in a computer graphics display (for instance a skybox) will produce identical images on the two retina (and hence on the two displays), imagery on flat frontal surfaces in 3D worlds (for example virtual white boards) produces nearly identical images, and 2D media are often presented on stereoscopic displays (including VR headsets). Our stimuli are directly comparable to these situations and predict increased tolerance to distortions that are binocularly unmatched and when the content is presented at a disparity relative the screen plane. In a VR headset there is no screen plane but the content is typically designed around an optimal working distance. As discussed in Section 6.1, further work is needed to determine whether the critical factor underlying the disparity pedestal effect seen here is the stereoscopic distance from the plane of best focus or some other parameter.

The difference in working distance of the display between a S3D screen and an HMD is one example of the influence of presentation mode. Another important factor is the independence of the left and right channels. A variety of formats for stereoscopic images and video exist including independent channels and displays (as in many HMDs or two-channel stereoscopes) and combined channels that present the image on the same physical display or in the same image stream. The left and right images can be combined in a channel in different ways: separated temporally on alternate frames; separated spatially on different parts of a combined image such as top versus bottom, left versus right, or odd versus even lines; or separated by color channel as in anaglyph displays. When such combined displays are compressed, artifacts are likely to differ in the left and right images even if the images are identical due to spatial and temporal dependencies in typical compression codecs. However, with completely independent channels that are processed identically the resulting left and right images should also be identical in cases where the source images are identical (as in the current study). In this instance, artifacts will also be identical in the two eyes images and likely more visible than if they were uncorrelated. Under these circumstances it might be

advisable to adjust the codec to decorrelate the artifacts in the two images. However, in practice most S3D content will consist of different left and right images due to binocular disparity and this will produce differences in the coding errors. Whether attempts to further decorrelate the images would improve performance is an open question. However, we did find one case where unmatched errors were more detectable than matched errors. In this instance there were particularly large interocular differences in distortions (Figure 10) which suggests that there might be limits to the ability to hide compression errors by binocular decorrelation.

7 | CONCLUSION


The main objectives of this study were to determine the effects of stereoscopic depth and the similarity of artifacts in the half images on artifact visibility. We found that for two specific images there was a significant interaction between the depth (disparity vs. no disparity) and compression (same or different) on the visibility of compression artifacts. The overall effect of depth on the perception of compression artifacts was significant and, for all the significant cases except Clipboard, compression artifacts were less visible for images with disparity than no disparity (i.e., stereo vision was silencing the artifact perception for these images). In the significant cases for matched compression versus unmatched compression, artifacts were significantly less perceptible in images with different compression compared to images with same compression. The exception again was Clipboard where the codec performed significantly better for same compression than for different compression, which was likely due to the high degree of compression artifact mismatch in this image. In general, our results suggest that barely detectable compression artifacts should be less problematic in stereoscopic 3D displays compared to non-stereoscopic displays at similar compression levels. Currently, the most ubiquitous use case for S3D content is VR and AR. The testing protocol presented here paves the way for development of effective AR/VR image quality test procedures which are essential to both codec development and to understanding the impact of compression on the AR/VR user experience.

ACKNOWLEDGEMENT

The authors would like to thank Bob Hou, Natan Jacobson and James Goel for their technical support.

ORCID

Sanjida Sharmin Mohona  <https://orcid.org/0000-0001-9091-9018>

Laurie M. Wilcox  <https://orcid.org/0000-0002-3594-6192>

Robert S. Allison  <https://orcid.org/0000-0002-4485-2665>

ENDNOTES

* We also used a small disparity range to minimize vergence-accommodation conflicts.³¹

† We deliberately chose this form of dichoptic imagery to isolate the effects of disparity relative to the screen and the difference in coding errors in the two images from the potential confounding effect of stereoscopic depth within the images (which would covary with our main variables).

REFERENCES

- Walls FG, MacInnis AS. VESA display stream compression for television and cinema applications. *IEEE J Emer Sel Top Circui Syst.* 2016;6(4):460–470. <https://doi.org/10.1109/JETCAS.2016.2602009>
- Descampe A, Keinert J, Richter T, Föföel S, Rouvroy G. JPEG XS, a new standard for visually lossless low-latency lightweight image compression. *Applications of Digital Image Processing XL.* Sep. 2017, vol. 10396;103960M. <https://doi.org/10.1117/12.2273625>
- Cai Q, Song L, Li G, Ling N. Lossy and lossless intra coding performance evaluation: HEVC, H.264/AVC, JPEG 2000 and JPEG LS. In: *Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference*; Dec. 2012. p. 1–9.
- “ST 2042–1:2012 - SMPTE Standard - VC-2 Video Compression,” *ST 2042–12012*, pp. 1–137, Aug. 2012, doi: 10.5594/SMPTE.ST2042-1.2012.
- ISO/IEC 21122–1:2019, “Information technology—JPEG XS low-latency lightweight image coding system—Part 1: Core coding system,” *ISO*. <http://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/07/45/74535.html> Accessed 29 Sept 2019.
- Walls F, MacInnis A. 27.4L: late-news paper: VESA display stream compression: an overview. *SID Symp Dig Tech Pap.* 2014;45(1):360–363. [10.1002/j.2168-0159.2014.tb00097.x](https://doi.org/10.1002/j.2168-0159.2014.tb00097.x)
- Video Electronics Standards Association. VESA display compression-M (VDC-M) standard, v 1.2. San Jose, USA; 2019.
- ISO. “ISO/IEC 29170-2.” International Organization of Standards. Switzerland: Geneva; 2015.
- Shahid M, Rossholm A, Lövsström B, Zepernick H-J. No-reference image and video quality assessment: a classification and review of recent approaches. *EURASIP J Image Video Process.* 2014;(1):40. <http://doi.org/10.1186/1687-5281-2014-40>
- Banks MS, Read JC, Allison RS, Watt SJ. Stereoscopic and the human visual system. *SMPTE Motion Imaging J.* 2012;121(4): 24–43.
- Au D, Mohona SS, Cutone MD, Hou Y, Goel J, Jacobson N, Allison RS, Wilcox LM. 3–4: stereoscopic image quality assessment. *SID Symp Dig Tech Pap.* 2019;50(1):13–16. <https://doi.org/10.1002/sdtp.12843>
- He L, Gao F, Hou W, Hao L. Objective image quality assessment: a survey. *Int J Comput Math.* 2014;91(11):2374–2388. <https://doi.org/10.1080/00207160.2013.816415>

13. Ouni S, Chambah M, Herbin M, Zagrouba E. Are existing procedures enough? Image and video quality assessment: review of subjective and objective metrics. *Image Quality and System Performance V*. 2008;6808:68080Q. <https://doi.org/10.1117/12.760803>
14. Chikkerur S, Sundaram V, Reisslein M, Karam LJ. Objective video quality assessment methods: a classification, review, and performance comparison. *IEEE Trans Broadc*. 2011;57(2):165–182. <https://doi.org/10.1109/TBC.2011.2104671>
15. Hoffman DM, Stoltzka D. A new standard method of subjective assessment of barely visible image artifacts and a new public database. *J Soc Inf Disp*. 2014;22(12):631–643.
16. Chandler DM. Seven challenges in image quality assessment: past, present, and future research. *ISRN Signal Process*, Vol. 2013;2013.
17. Brunnström K, Allison RS, Chandler DM, Colett H, Corriveau P, Daly S, Goel J, Knopf J, Wilcox LM, Yaacob Y, Yang SN. Industry and business perspectives on the distinctions between visually lossless and lossy video quality: mobile and large format displays. *Electron Imaging*. 2017;2017(14):118–133. <https://doi.org/10.2352/ISSN.2470-1173.2017.14.HVEI-131>
18. Larson EC, Chandler DM. Most apparent distortion: full-reference image quality assessment and the role of strategy. *J Electron Imaging*. Jan. 2010;19(1):011006. <https://doi.org/10.1117/1.3267105>
19. Allison RS, Wilcox LM, Wang W, Hoffman DM, Hou Y, Goel J, Deas L, Stoltzka D. 75–2: invited paper: large scale subjective evaluation of display stream compression. *SID Symp Dig Tech Pap*. 2017;48(1):1101–1104. <https://doi.org/10.1002/sdtp.11838>
20. Suchow JW, Alvarez GA. Motion silences awareness of visual change. *Curr Biol*. 2011;21(2):140–143. <https://doi.org/10.1016/j.cub.2010.12.019>
21. Choi LK, Cormack LK, Bovik AC. Motion silencing of flicker distortions on naturalistic videos. *Signal Process Image Commun*. 2015;39:328–341. <https://doi.org/10.1016/j.image.2015.03.006>
22. Chen MJ, Bovik AC, Cormack LK. Study on distortion conspicuity in stereoscopically viewed 3D images. 2011 IEEE 10th IVMSP Work: Percep Visual sig Anal. 2011;24–29. <https://doi.org/10.1109/IVMSPW.2011.5970349>
23. Chen M-J, Cormack LK, Bovik AC. Distortion conspicuity on stereoscopically viewed 3D images may correlate to scene content and distortion type. *J Soc Inf Disp*. 2013;21(11):491–503. <https://doi.org/10.1002/jsid.198>
24. Shao F, Lin W, Gu S, Jiang G, Srikanthan T. Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics. *IEEE Trans Image Process*. 2013;22(5):1940–1953. <https://doi.org/10.1109/TIP.2013.2240003>
25. Zhou W, Yu L. Binocular responses for no-reference 3D image quality assessment. *IEEE Trans Multimed*. 2016;18(6):1077–1084. <https://doi.org/10.1109/TMM.2016.2542580>
26. Liu Y, Yang J, Meng Q, Lv Z, Song Z, Gao Z. Stereoscopic image quality assessment method based on binocular combination saliency model. *Signal Process*. 2016;125:237–248. <https://doi.org/10.1016/j.sigpro.2016.01.019>
27. Campisi P, Callet PL, Marini E. Stereoscopic images quality assessment. In: 2007 15th European Signal Processing Conference; Sep. 2007. p. 2110–2114.
28. Seuntjens P, Meesters L, Ijsselstein W. Perceived quality of compressed stereoscopic images: effects of symmetric and asymmetric JPEG coding and camera separation. *ACM Trans Appl Percept TAP*. 2006;3(2):95–109.
29. Wang X, Yu M, Yang Y, Jiang G. Research on subjective stereoscopic image quality assessment. *Multim Content Access: Algo Sys III*. 2009;7255:725509.
30. Yang J, Hou C, Zhou Y, Zhang Z, Guo J. Objective quality assessment method of stereo images. In: 3DTV Conference: The true vision-capture, transmission and display of 3D video, 2009; 2009. p. 1–4.
31. Chen C, Deng H, Zhong F, Ji Q, Li Q. Effect of viewpoints on the accommodation response in integral imaging 3D display. *IEEE Photonics J*. 2020;12(3):1–14. <https://doi.org/10.1109/JPHOT.2020.2993575>
32. Agresti A. Random effects: generalized linear mixed models. In: *An introduction to categorical data analysis*; 2007.
33. Cutone MD, Dalecki M, Goel J, Wilcox LM, Allison RS. P-31: a statistical paradigm for assessment of subjective image quality results. *SID Symp Dig Tech Pap*. 2018;49(1):1312–1314. <https://doi.org/10.1002/sdtp.12154>
34. Wilkinson GN, Rogers CE. Symbolic description of factorial models for analysis of variance. *J R Stat Soc Ser C Appl Stat*. 1973;22(3):392–399. <https://doi.org/10.2307/2346786>
35. Cook GW, Ribera J, Stoltzka D, Xiong W. 85-4: a subpixel-based objective image quality metric with application to visually lossless image compression evaluation. *SID Symp Dig Tech Pap*. 2018;49(1):1163–1166. <https://doi.org/10.1002/sdtp.12103>
36. Ribera J, Cook GW, Stoltzka D, Xiong W. 68-1: a machine learning approach to objective image quality evaluation. *SID Symp Dig Tech Pap*. 2019;50(1):957–960. <https://doi.org/10.1002/sdtp.13084>
37. Mohona SS, Au D, Kio OG, Robinson R, Hou Y, Wilcox LM, Allison RS. Subjective assessment of stereoscopic image quality: the impact of visually lossless compression. In: 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX); May 2020. p. 1–6. [10.1109/QoMEX.48832.2020.9123129](https://doi.org/10.1109/QoMEX.48832.2020.9123129).
38. Regan D, Gray R. Binocular processing of motion: some unresolved questions. *Spat Vis*. 2009;22(1):1–43. <https://doi.org/10.1163/156856809786618501>
39. Zaroff CM, Knutelska M, Frumkes TE. Variation in stereoacuity: normative description, fixation disparity, and the roles of aging and gender. *Invest Ophthalmol Vis Sci*. Feb. 2003;44(2):891–900. <https://doi.org/10.1167/iovs.02-0361>
40. Maddox EE. Investigations in the relation between convergence and accommodation of the eyes. *J Anat Physiol*. 1886;20(Pt 3):475–508.
41. Hoffman DM, Girshick AR, Akeley K, Banks MS. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *J Vis*. 2008;8(3):33–33, 3330. <https://doi.org/10.1167/8.3.33>
42. Yang S, Sheedy JE. Effects of vergence and accommodative responses on viewer's comfort in viewing 3D stimuli. *Stereosc Displays Appl XXII*. 2011;7863:78630Q. <https://doi.org/10.1117/12.872546>
43. Marcos S, Moreno E, Navarro R. The depth-of-field of the human eye from objective and subjective measurements.

- Vision Res. 1999;39(12):2039–2349. [https://doi.org/10.1016/S0042-6989\(98\)00317-4](https://doi.org/10.1016/S0042-6989(98)00317-4)
44. Schor C. Fixation disparity: a steady state error of disparity-induced vergence. *Am J Optom Physiol Opt.* 1980;57(9): 618–631.
 45. Legge GE. Binocular contrast summation—I. Detection and discrimination. *Vision Res.* 1984;24(4):373–383. [https://doi.org/10.1016/0042-6989\(84\)90063-4](https://doi.org/10.1016/0042-6989(84)90063-4)
 46. Howard IP, Rogers BJ. Binocular summation, masking, and transfer. In: *Perceiving in depth. 2 Stereoscopic vision*, Oxford: Oxford University Press; 2012. p. 107–147.
 47. Baker DH, Meese TS. Binocular contrast interactions: dichoptic masking is not a single process. *Vision Res.* 2007;47(24): 3096–3107. <https://doi.org/10.1016/j.visres.2007.08.013>
 48. Poirson AB, Wandell BA. Pattern-color separable pathways predict sensitivity to simple colored patterns. *Ophthalmic Lit.* 1997;1(50):50.
 49. Wichmann FA, Hill NJ. The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys.* Nov. 2001; 63(8):1293–1313. <https://doi.org/10.3758/BF03194544>
 50. McKee SP, Klein SA, Teller DY. Statistical properties of forced-choice psychometric functions: implications of probit analysis. *Percept Psychophys.* 1985;37(4):286–298. <https://doi.org/10.3758/BF03211350>

AUTHOR BIOGRAPHIES



Sanjida Sharmin Mohona received her BSc degree in electronics and communication engineering from Khulna University of Engineering and Technology, Bangladesh, in 2014 and MSc degree in computer engineering from York University, Canada, in 2018, and is pursuing a PhD degree in computer engineering at York University, Canada.



Laurie M. Wilcox obtained her PhD in Psychology from Western University in 1992. She is a Professor of Psychology and member of the Centre for Vision Research at York University. She is also appointed to the graduate program in Biology. She was the recipient of a NSERC Women's Faculty award which supported her appointment to York University in 1996. Her research focusses on both basic and applied aspects of 3D vision including virtual reality,

3D film, and 3D display systems. Her research is supported by the Natural Sciences and Engineering Council of Canada and the Ontario Centres of Excellence. She is a fellow of the Association of Psychological Science, serves on the editorial board for the *Journal of Perceptual Imaging*, is President-Elect for the Vision Sciences Society, and is a member of SPIE and IEEE.



Robert Allison is a Professor of Electrical Engineering and Computer Science at York University and a member of the Centre for Vision Research. He is also appointed to the graduate program in Psychology at York. He received his BSc in Computer Engineering from the University of Waterloo in 1991. After graduation he worked as an Electrical Engineer at Atlantis Aerospace in Brampton, Canada, where he designed electronics for flight training devices. Following this he completed a MSc in Electrical Engineering (Biomedical Engineering) from the University of Toronto before obtaining his PhD specializing in stereoscopic vision from York University in 1998. He was on the experimental team for the 1998 Neurolab space shuttle mission and did post-doctoral research at York University and the University of Oxford. His research enables effective technology for advanced virtual reality and augmented reality and for the design of stereoscopic displays. He is a recipient of the Premier's Research Excellence Award and a York Research Chair in recognition of this work. He is a Senior Member of the Institute of Electrical and Electronic Engineers and also a member of the IEEE Computer Society and the Association for Computing Machinery.

How to cite this article: Mohona SS, Wilcox LM, Allison RS. Subjective assessment of display stream compression for stereoscopic imagery. *J Soc Inf Display.* 2021;29:591–607. <https://doi.org/10.1002/jsid.1002>