Luminance Asymmetry in Stereoscopic Content: Binocular Rivalry or Luster

Marja Salmimaa[1], Jussi Hakala[2], Monika Pölönen[1], Toni Järvenpää[1], Radu Bilcu[1], Jukka Häkkinen[2]
[1]Media Technologies Laboratory, Nokia Research Center, Finland
[2]Department of Media Technology, Aalto University, Finland

Abstract
Binocular representation of images suffers often from asymmetries between the images, originating from the display systems used for the presentation and from mismatches in the left-right images introduced during the image acquisition and processing. Two experiments were conducted to investigate the implications of the altered luminance levels of the left-eye and right-eye images to the visual experience.

1. Introduction
Stereo vision is the ability to perceive depth in the visual information received by the eyes and processed in the visual cortex of the brain. The perception of depth is a sophisticated process involving different depth cues, which are often categorized to monocular (pictorial), oculomotor and binocular depth cues. Numerous monocular depth cues (see [13] for reviews) include e.g. interposition, linear perspective, light and shade, relative size and texture gradient (see e.g. [5]), aerial perspective, and motion parallax [8], [21]. Oculomotor cues are based on feedback from the muscles used to control the vergence movements and accommodation of the eye [7], [20].

1.1. Binocular vision
Binocular depth cues are originating from the lateral location of the eyes, which are a small distance apart from each other. Because of the difference in the horizontal position of the eyes, the retinal images are formed from slightly different vantage points. Consequently, depth differences produce differences to the retinal images of the left and right eye. The differences in the received information are processed in the brain, where the fusional mechanism combines these two retinal images into one image inducing sensation of the depth in the content. The depth sensation arising from the binocular disparity is called stereopsis. Stereopsis was first explained by Charles Wheatstone in 1838 and later by Helmholtz and Julesz [26], [8], [12]. A review of human stereopsis is provided by Patterson and Martin [17].

In addition to stereopsis, binocular rivalry is a central phenomenon in binocular vision. In rivalry, one monocular view dominates at a time and the perception alternates between different images. Furthermore, stereopsis and binocular rivalry may coexist [27]. Binocular rivalry may arise between any images if they differ sufficiently. These differences may occur e.g. in colors (binocular color rivalry) or object outlines (binocular contour rivalry). When the stimuli differ in the luminances, rivalry is not the only possible perceptual outcome. Depending on the stimulus features, binocular summation [2], binocular luster [18], the sieve effect or the floating effect might occur [9], [6]. Furthermore Kumar found that the local regions of the images may be perceived as slanted if the luminance presented for each of the eyes differ in that region [14].

Determining the threshold values for observable or acceptable luminance disparity, or the threshold reducing the viewing comfort is not straightforward. Earlier research exists, however some of the related documents do not disclose references or supporting research results. Also systematic studies with natural content are relatively rare. In earlier studies Beldie reports observable overall luminance asymmetries differences of 3-6 dB, and 0.2 dB for moving and sharply localized luminance deviations [3]. In Self's review [24], recommendations for less than 50% (preferably less than 25 %) and smaller than 10% luminance difference can be found. ISO 9241-303 states the interocular luminance difference for binocular displays should not exceed 10% [11]. The differences in the recommendations probably come from differences in contents as well as the evaluation criteria.

1.2. Binocular systems
Binocular display systems provide a binocular, possibly stereoscopic representation of the content for the observer. Examples of such systems are head-mounted displays (HMD) (see [15], [4] for reviews). In addition 3D displays can be considered as binocular display systems, as those displays, either glasses-based or autostereoscopic, separate the left- and the right-eye information for proper presentation of the spatial shift or binocular disparity in the content (see [16] for a review).

However, with the above-mentioned display systems the resulting visual experience is in addition to the display specific parameters depending on the capture and processing of the information rendered on the display(s). Stereo image pairs are often captured with devices equipped with two camera sensors. Each sensor will capture one image in raw format, and the images are processed mostly separately in the image processing chain. Some mismatch between the left and right images appears already at sensor level, due to the manufacturing tolerances. E.g. the effective sensitivity of the imaging sensors may differ. As a consequence, even if the same sensor settings would be used for the capturing, the color and the luminance of the two images can be different. The problem becomes even more evident when the lightning of the two camera views is different due to the scene geometry. Furthermore, several processing steps in the image processing chain (e.g. auto-white balance, tone mapping to mention some) potentially introduce mismatches in the two images. These processing steps may occur in several places in the stereo capturing system. Firstly, the imaging sensors can be driven by different settings (exposure time and analog gain as examples) to reduce the influence of the different sensor sensitivity. Secondly the stereo capturing system may contain a post processing step called color matching, trying to eliminate the color and luminance asymmetry introduced by the whole image processing chain.

One important aspect that needs to be addressed also in the context of stereo imaging, is the efficient encoding of the captured stereo image pair. An interesting approach, which appeared in the context of video coding but can be utilized also for efficient image coding, is asymmetric coding. In such a scenario, in order to increase the coding efficiency, one image might be coded with full color and luminance, while the second image can be encoded with reduced color and luminance, or by using some methods reducing the luminance of the processed image [1]. However, understanding the significance of the luminance asymmetry to the depth perception and the overall visual experience is needed before applying these coding schemes.

In this paper the implications of the luminance difference between the left-eye and right-eye images to the visual experience were investigated in two experiments. The next section introduces the first experiment where the acceptability of the images was determined in three test conditions: (a) luminance of the image displayed to one eye was kept constant when the other image's luminance was decreased, (b) the luminance of both images was decreased symmetrically, and (c) the luminance of the image displayed to one eye was decreased and for the other image increased, keeping the overall luminance constant. Section 3 provides the results for the experiment 1. Section 4 describes the experiment 2 where the overall viewing experience, image naturalness and depth impression were assessed. Section 5 presents the results of the second experiment and finally, Section 6 concludes the paper.

2. Methods - experiment 1

2.1. Equipment

As crosstalk noticeably degrades the quality of the stereoscopic images, and may influence on the depth perception both from disparity and monocular depth cues [25], experiments described in this paper were conducted by using a stereoscope providing zero crosstalk condition i.e. complete channel separation. Furthermore, this way the experiments were more controlled and the results better applicable in general in different display domains.

The stimuli were displayed on a 22.2" ViewSonic VP2290b display with a resolution of 204 pixels per inch on a gray background. The optical properties for the display were determined by measuring grey levels with a spectroradiometer (PhotoResearch SpectraScan PR-670). The luminance was measured through the stereoscope at the viewing distance (40 cm) at nine different grey levels. Based on the measurements the appropriate image RGB levels were found to produce the desired luminance reductions. The obtained luminance values for grey levels 255 (white) and 0 (black) were 126.6 cd/m2 and 0.4cd/ m2, respectively.

2.2. Procedure

The experiment had three phases and each phase was done on a different day. The order of the phases for each participant was counterbalanced. In each phase the stimuli consisted of the same five stereoscopic images of natural scenes selected carefully avoiding excessive disparities. The participants viewed the images through a fixed mirror stereoscope in front of the display and the optical distance from the participant's eye to the display plane was 40 cm. Distance to the display plane was controlled by a chin and forehead rest.

The display duration of the stimuli was limited only by the participant's response, which triggered the display of the next stimulus. A fixation cross was displayed for one second between the stimuli. The image pair's luminance was changed based on the participants' responses and the current phase of the experiment. In phase 1 the image displayed to the other eye was always the original image, and the other eye's image luminance was decreased. In phase 2 the luminance of both images was decreased symmetrically. In phase 3 the other image's luminance was decreased and the other's increased, keeping the overall luminance constant. In phase 3 the starting luminance for both images was thus 50% of the original.

In more detail, the experimental design was a parallel staircase experiment with ten concurrent staircases, each eye-scene combination having their own staircase. The order of the stimuli in each staircase was randomized. The participants were instructed to press 'y' or 'n' on the computer keyboard to indicate whether the stereoscopic image was acceptable or not. For the next stimuli in the staircase the luminance was decreased by one decibel if the participant pressed 'y', increased by one decibel if the participant pressed 'n', or kept unaltered if the participant pressed 'n' but the luminance was already at the original level. The termination criterion for each staircase was ten reversals, not including the first reversal. When a staircase was terminated, the other staircases still continued until all of them had reached the termination criterion.

2.3. Stimuli

The stimuli were selected to represent camera phone clusters in photo space described by International Imaging Industry Association I3A for the objective characterization of camera attributes and subjective image quality experiments [10]. I3A's White paper identifies six clusters, and originally corresponding six stimuli was prepared. However, based on the expert evaluations and a pilot test, from the original six test stimuli only five qualified for the experiments. One of the test images had some rendering artefacts potentially affecting the results, thus it was discarded. Following stimuli were used: Image 1 'Bus stop' (cluster 6, average luminance 32.8 cd/m²), Image 2 'Woman and man with a cast' (cluster 2, average luminance 29.3 cd/m²), Image 3 'Graffiti building and boys'(cluster 3, average luminance 3.9 cd/m²), Image 4 'Boy in the forest' (cluster 1, average luminance 2.0 cd/m²), Image 5 'Boy and girl sitting'(cluster 5, average luminance 9.9 cd/m²). Each eye's image was 800 pixels wide and 480 pixels high. The stimuli were displayed side by side on the display on a gray background.

2.4. Participants

Twelve participants took part in the experiment. Participants were required to have normal or corrected to normal far and near visual acuity (Lea-numbers $\geq$0.50), and normal stereo acuity $\leq$ 240 secs-of-arc (plate V; TNO test plates V-VII), interpupillary distance between 51-71 mm, near horizontal phoria not more than 13D exophoria or 7D esophoria, and vertical phoria 1D at the maximum (Maddox Wing test). Near point of convergence (NPC) and near point of accommodation (NPA) were also examined but had no exclusion criteria. Four participants wore glasses.

Ages of the participants were between 20 and 33 years with a mean of 26.0, five participants were women and seven men. Participants were recruited from the staff and students of the university and they were not rewarded for their participation. One participant completed two of the three phases after which she reported having migraine like symptoms and she was instructed not to complete the third phase.

3. Results from the experiment 1

A total of 10806 evaluations were made during the three phases. Due to different evaluation criteria some participants' data were excluded from the analysis. In phase 1 four of the participants had still found acceptable a condition where the other eye's image was completely black, thus their data were excluded from further analysis. In phase 2 three participants had used a much lower criterion than the rest. They had found all the luminance levels displayed acceptable, and thus their data were excluded. In phase 3 we found two nearly equal sized groups (group 1, N=5, group 2, N=6) that had a significantly different threshold than the other. A probit analysis was performed on the data to find the psychometric curves and acceptability thresholds. The 75% acceptability thresholds for the lowered luminance, in percentages of the original, were found to be at 40, 44 and 28 percent in phases 1-3 respectively. In phase 1 this means a luminance reduction of 60% in one image, and in phase 2 the total luminance reduction threshold is 56%. In phase 3, where the other image's luminance was increased while the other image's luminance decreased, the luminance reduction threshold from the higher image luminance to the lower image luminance is 61%, which verifies the result of phase 1.

4. Methods - experiment 2

In the experiment 2 same equipment and stimuli than in the experiment 1 were used. Luminance levels for the second experiment were obtained based on the acceptability threshold found in the first experiment. Both higher and lower luminance levels compared to the acceptability threshold were used and the luminance was altered in 2dB increments.

4.1. Procedure

The experiment setup was the same as in Experiment 1. However in this experiment the participants evaluated the stimuli on three Likert-like scales with discreet numbers from 1-9 on the computer keyboard. The scales were viewing experience, naturalness and depth impression. Each scene-eye-luminance combination was repeated in a randomized order three times in different blocks.

4.2. Participants

Eighteen participants participated in the experiment, each had normal or corrected to normal visual acuity (Lea-numbers ≥0.50), and normal stereo acuity ≤ 240 secs-of-arc (plate V; TNO test plates V-VII), interpupillary distance between 51-71 mm, near horizontal phoria not more than 13D exophoria or 7D esophoria, and vertical phoria 1D at the maximum (Maddox Wing test). Near point of convergence (NPC) and near point of accommodation (NPA) were also examined but had no exclusion criteria. Four participants wore glasses. Ages of the participants were between 20 and 36 years with a mean of 25.3, seven participants were women and eleven men. All participants were staff or students of the university and they were rewarded with a lunch voucher for their participation.

5. Results from the experiment 2

Because of the nature of the data, non-parametric test procedures were used for statistical analysis. A Holm-Bonferroni correction was used to control for the occurrence of false positive p-values.

5.1. Luminance differences

The Kruskal–Wallis test revealed significant differences between overall viewing experience (χ^2=485.933, df=5, p<0.000), naturalness (χ^2=313.543, df=5, p<0.000), and perceived depth (χ^2=310.061, df=5, p<0.000) scores. Pair comparison of the data showed that almost all the luminance pairs differed significantly in all the three variables evaluated. As Figure 1 shows, overall viewing experience, image naturalness, and perceived depth decreased as the luminance difference between the eyes increased.

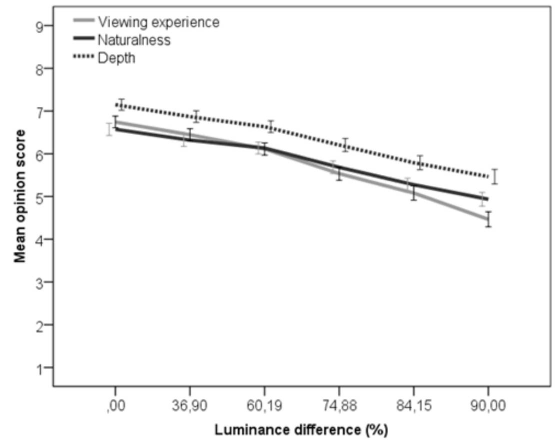

Figure 1. Mean opinion scores (MOS) for luminance differences. Vertical lines represent error bars, confidence interval level 95%.

Also Kendall's correlation analysis revealed small, but significant negative correlations between luminance differences and viewing experience (r=-0.302, p<0.01), luminance differences and naturalness (r=-0.242, p<0.01), and luminance differences and perceived depth (r=-0.242, p<0.01). As expected on the basis of previously published results, perceived depth impression was less sensitive to the changes in luminance levels than other evaluated parameters [19].

5.2. Content differences

The Kruskal–Wallis test revealed significant differences between the contents in viewing experience (χ^2=104.720, df=4, p<0.000), naturalness (χ^2=130.447, df=4, p<0.000), and perceived depth (χ^2=165.031, df=4, p<0.000) scores. Pair comparison of contents showed that most of the scores given to the overall image quality, naturalness, and depth impression were significantly different.

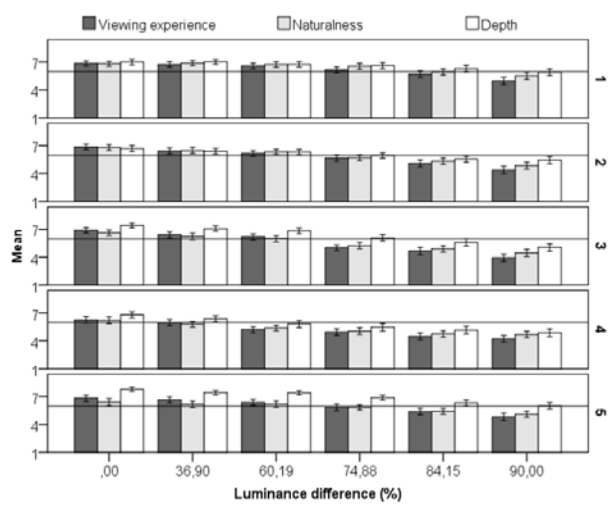

Figure 2. MOS for different contents with different luminance combinations. Reference line is the overall mean value (5.97) and vertical lines represent error bars, confidence interval level 95%.

As Figure 2 shows, some of the contents, such as images 1 and 5, were scored relatively high even when luminance difference of 84,15% was used, whereas overall viewing experience and naturalness of image 4 decreased already when luminance difference of 60,19% was used. Thus it can be assumed that content specific parameters affect image quality evaluations, but because only five different contents of natural images were used in our tests, it is difficult to say which content specific parameters are more sensitive to the changes in luminance levels. For reliable generalizations, different contents with controlled content-specific features should be studied, see e.g. [22], [23].

6. Conclusions

The implications of the altered luminance levels of the left-eye and right-eye images to the visual experience were investigated. First the acceptability of the images was determined in three test conditions. Secondly the overall viewing experience, image naturalness and depth impression were assessed. Participants found surprisingly large binocular luminance differences acceptable. This might have occurred because luminance differences do not always produce binocular rivalry, but luster effect occurs. As binocular luster is a common occurrence in the natural world [6], it might be that luminance differences in natural images are regarded as surface glossiness rather than unnatural stimuli. Furthermore, as expected, content specific features were important parameters when image naturalness, overall quality, and perceived depth were evaluated. Images with dark background or low target-background contrast were more sensitive to luminance changes than images with brighter backgrounds and clear contrast differences.

Evaluated levels of perceived depth were less sensitive to luminance changes than overall image quality and naturalness; increase in luminance differences between the eyes decreased perceived depth impressions scores similar to the changes in image naturalness and overall quality. However scores given to the depth impression were always higher and the influence of luminance asymmetry varied between contents; with some contents' impression of depth remained good regardless of the scores of overall quality and naturalness.

In practice obtained results may give guidance e.g. for the asymmetric image coding schemes and may be used when developing methodologies and assessing parameters for the image processing pipeline in the stereo imaging.

7. References

[1] P. Aflaki, M. M. Hannuksela, J. Hakala, J. Häkkinen, M. Gabbouj, "Joint adaptation of spatial resolution and sample value quantization for asymmetric stereoscopic video compression: a subjective study," Proc. International Symposium on Image and Signal Processing and Analysis (ISPA), 396-401 (2011).

[2] S.M. Anstis, A. Ho, "Nonlinear combination of luminance excursions during flicker, simultaneous contrast, afterimages and binocular fusion," Vision Research **38**, 523–539 (1998).

[3] I. Beldie, "Luminance asymmetry in stereo TV images," Proceedings of SPIE **1457**, 242-247 (1991).

[4] O. Cakmakci, J. Rolland, "Head-Worn Displays: A Review," Journal of Display Technology **2**, 199-216 (2006).

[5] J.E. Cutting, R.T. Millard, "Three gradients and the perception of flat and curved surfaces," Journal of experimental psychology: General **113**, 198-216 (1984).

[6] M. A. Formankiewicz, J. D. Mollon, "The psychophysics of detecting binocular discrepancies of luminance," Vision Research **49**, 1929-1938 (2009).

[7] W. C. Gogek, "Convergence as a cue to absolute distance," Report No. 467. U. S. Army Medical Research Laboratory, Fort Knox, Kentucky, USA (1961).

[8] H. Helmholtz, "Treatise on physiological optics," (J.P.C. Southall, Trans.) Dover, New York, NY (1924). Original work published 1909.

[9] I.P. Howard, "Depth from binocular rivalry without spatial disparity," Perception **24**, 67–74 (1995).

[10] I3A International Imaging Industry Association, "The Camera Phone Image Quality (CPIQ) Initiative White Paper – Fundamentals and review of considered test methods," (2007). http://www.i3a.org/resources/

[11] ISO 9241-303, "Ergonomics of human-system interaction, Part 303: Requirements for electronic visual displays Annex E; Virtual display – Performance objectives," International Standard (2008).

[12] B. Julesz, "Foundations of cyclopean perception," University of Chigago Press, Chigago, IL (1971).

[13] L. Kaufman, "Sight and mind: An introduction to visual perception," Oxford University Press, Oxford, England (1974).

[14] T. Kumar, "Stereopsis due to luminance differences in the two eyes," Vision research **35**, 255-262 (1995).

[15] J.E. Melzer, K. Moffit, "Head-Mounted Displays designing for the user" McGraw-Hill USA (1997).

[16] S. Pastoor, M. Wöpking, "3-D displays: A review of current technologies," Displays **17(2)**, 100-110 (1997).

[17] R. Patterson, W.L. Martin, "Human stereopsis," Human Factors **34(6)**, 669-692 (1992).

[18] W. Pieper, I. Ludwig, "The minimum luminance-contrast requirements for stereoscopic lustre," Perception **31** ECVP Abstract Supplement, 185 (2002).

[19] M. Pölönen, M. Salmimaa, J. Häkkinen, "Effect of ambient illumination level on perceived autostereoscopic display quality and depth perception," Displays **32(3)**, 135–141 (2011).

[20] W. Richards, J.F. Miller, "Convergence as a cue to depth," Perception and Psychophysics **5**, 317-320 (1969).

[21] B. Rogers, M. Graham, "Motion parallax as an independent cue for depth perception," Perception **8**, 125–134 (1979).

[22] M. Salmimaa, T. Järvenpää, M. Pölönen, "2D viewing experience with fixed 3D displays," Proc. SPIE **7863**, 78631R (2011).

[23] M. Salmimaa, T. Järvenpää, M. Pölönen, "Viewing experiences of two-dimensional photos and textual content on autostereoscopic 3-D displays," Journal of Electronic Imaging **21**, 011010 (2012).

[24] H. Self, "Optical tolerances for alignment and image differences for binocular helmet-mounted displays," Technical Report AAMRL-TR-86-019, Harry G. Armstrong Aerospace Medical Research Lab, Wright-Patterson AFB, USA (1986).

[25] I. Tsirlin, L.M. Wilcox, R.S. Allison, "The Effect of Crosstalk on the Perceived Depth From Disparity and Monocular Occlusions," IEEE Transactions on Broadcasting **57**, 445-453 (2011).

[26] C. Wheatstone, "Contributions to the physiology of vision.—Part the First. On some remarkable, and hitherto unobserved, phænomena of binocular vision," Philosophical Transactions of the Royal Society of London **128**, 371-394 (1838).

[27] J.M. Wolfe, "Stereopsis and Binocular Rivalry," Psychological Review **93**, 269-28 (1986).